

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2004, 20:34:22 ; Search time 6775 Seconds
(without alignments)
7682.624 Million cell updates/sec

Title: US-10-069-541-5
Perfect score: 1743
Sequence: 1 atggctttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*

28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query					Description
	No.	Score	Match Length	DB	ID	
	1	1473	84.5	1743	29	AY413298 Homo sapi
	2	1463	83.9	1743	29	AY413299 Pan trogl
	3	1375	78.9	4097	11	AK053063 Mus muscu
	4	1375	78.9	4306	11	AK034415 Mus muscu
	5	1156.2	66.3	1743	29	AY413300 Mus muscu
	6	518.8	29.8	672	29	AG157499 Pan trogl
	7	472.2	27.1	707	14	CD350164 UI-M-FY0-
	8	462.8	26.6	669	13	BY727598 BY727598
	9	404	23.2	516	10	BE233479 139685 MA
	10	329.8	18.9	650	10	BB626260 BB626260
	11	312.8	17.9	541	10	AW668962 111664 MA
	12	290	16.6	675	13	BY729567 BY729567
	13	274.2	15.7	524	10	BE723927 198406 MA
	14	263.2	15.1	800	9	AL669749 AL669749
	15	225.8	13.0	549	13	BW274870 BW274870
	16	212.4	12.2	1037	9	AL666817 AL666817
	17	210	12.0	941	14	CD360297 AGENCOURT
	18	209.2	12.0	641	12	BI630566 RH59836.5
	19	207.4	11.9	640	12	BI629504 RH58381.5
	20	205.2	11.8	658	12	BM629925 170006875
	21	203.2	11.7	652	10	BB626456 BB626456
	22	193.6	11.1	583	13	BW277281 BW277281
	23	192.6	11.0	605	13	BQ829470 LL6in2176
	24	186.4	10.7	624	12	BJ122485 BJ122485
	25	183	10.5	681	14	CD306544 StrPu691.
	26	178.6	10.2	565	12	BJ125564 BJ125564
	27	177.4	10.2	576	14	CB391304 OSTF149A8
	28	176.6	10.1	310	9	AL918603 AL918603
	29	166	9.5	604	9	AU199794 AU199794
	30	163.4	9.4	500	9	AV994375 AV994375
	31	163.2	9.4	555	12	BJ117801 BJ117801
	32	158.4	9.1	646	9	AB078155 AB078155
c	33	157.4	9.0	801	13	BW002036 BW002036
c	34	155	8.9	561	28	AQ316435 RPCI11-10
	35	153	8.8	584	12	BJ105382 BJ105382
	36	145.2	8.3	500	12	BP186503 BP186503
c	37	134	7.7	632	28	AZ612750 1M0439J17
	38	128.2	7.4	500	12	BJ105730 BJ105730
c	39	127.6	7.3	618	28	AZ908709 RPCI-24-2
	40	121	6.9	926	29	AL295624 Tetraodon
c	41	117.2	6.7	525	12	BI508286 BB170004A
c	42	117.2	6.7	530	12	BI503332 BB170012A
	43	112.8	6.5	355	9	AU209671 AU209671
c	44	112.4	6.4	558	12	BI507950 BB170010A
c	45	108	6.2	420	12	BI506529 BB170027B

ALIGNMENTS

RESULT 1

AY413298

LOCUS AY413298 1743 bp DNA linear GSS 12-DEC-2003

DEFINITION Homo sapiens HCM4844 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY413298

VERSION AY413298.1 GI:39769260

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..1743
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene <1..>1743
 /locus_tag="HCM4844"

ORIGIN

Query Match 84.5%; Score 1473; DB 29; Length 1743;

Best Local Similarity 84.5%; Pred. No. 0;

Matches 1473; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

```

Qy      1 ATGGCTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
          |||
Db      1 ATGGCTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
          |||
Db     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

```

Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGNN	180
Qy	181	ACCTGGGTCCGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	NN	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	NN	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	NN	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	361	NN	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	NN	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020

Db	961	 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	 CGGAACATCTACCAGCTTTCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	 ATGCGAATCACAGTGTGTTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	 GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	 TGA	1743

RESULT 2

AY413299
 LOCUS AY413299 1743 bp DNA linear GSS 12-DEC-2003
 DEFINITION Pan troglodytes HCM4844 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
 ACCESSION AY413299
 VERSION AY413299.1 GI:39769261
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 REFERENCE 1 (bases 1 to 1743)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1743)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
 FEATURES
 Location/Qualifiers
 source 1..1743
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 gene <1..>1743
 /locus_tag="HCM4844"
 ORIGIN

Query Match 83.9%; Score 1463; DB 29; Length 1743;
 Best Local Similarity 84.1%; Pred. No. 0;
 Matches 1466; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

```

Qy      1 ATGGCTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
          |||
Db      1 ATGGCTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy      61 GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
          |||
Db      61 GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
          |||
Db      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACANN 180

Qy      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240

```

[illegible]

Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACGTCATCTTAACCAACATTTGCGTCTCCTATCTAGCCAAATAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAANGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 3

AK053063

LOCUS AK053063 4097 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched library, clone:D930038E20 product:solute carrier family 5 (choline transporter), member 7, full insert sequence.

ACCESSION AK053063

VERSION AK053063.1 GI:26343192
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6 (bases 1 to 4097)
 AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
 Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
 Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
 Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
 Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
 Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers
source 1. .4097
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D930038E20"
/db_xref="MGI:2424012"
/db_xref="taxon:10090"
/clone="D930038E20"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="15 days embryo"

CDS 512..2254
/note="unnamed protein product; putative solute carrier family 5 (choline transporter), member 7 (MGD|MGI:1927126, GB|NM_022025, evidence: BLASTN, 99%, match=1743)"
/codon_start=1
/protein_id="BAC35253.1"
/db_xref="GI:26343193"
/translation="MSFHVEGLVAIILFYLLIFLVGIWAAWKTKNSGNPEERSEAIIV GGRDIGLLVGGFTMTATWVGGGYINGTAEAVYGPGCGLAWAQAPIGYSLSLILGGLFF AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAFSALGATISVIIDVD VNISVIVSALIALIYLTLVGGLYSVAYTDVVQLFCIFIGLWISVPFALSHPA VTDIGFT AVHAKYQSPWLGTIESVEVYTWLDNFLLLMLGGIPWQAYFORVLSSSSATYA QVLSFL AAFGCLVMALPAICIGAIGASTDWNQTAYGYPDPKTKKEADMI LPIVLQYLCPVYISF FGLGA VSAAMSSADSSILSASSMFARNIYQLSFRQNASDK EIVWMRITVLVFGASA TAMALLTKTVYGLWYLSSDLVYIIIFPQLLCVLFIKGTNTYGA VAGYIFGLFLRITGG EPYLYLQPLIFYPGYYSDKNGIYNQRFPFKTLSMVSFFT NICVSYLA KYLFESGTL PPKLDVFDAVARHSEENMDKTILVRNENIKLNELAPVKPRQSLT LSSTFTNK EALLDV DSSPEGSGTEDNLQ"

ORIGIN

Query Match	78.9%;	Score 1375;	DB 11;	Length 4097;
Best Local Similarity	86.8%;	Pred. No. 0;		
Matches 1513;	Conservative	0;	Mismatches 230;	Indels 0; Gaps 0;

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG 60
        ||| ||||| | | ||||| | ||||| | | ||||| | | ||||| | | |||||
Db     512 ATGTCTTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTCTG 571
    
```

Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
 |||||
 Db 572 GTTGAATATGGGCTGCATGGAAAACCAAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 631

Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
 |||||
 Db 632 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC 691

Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
 |||||
 Db 692 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 751

Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
 ||
 Db 752 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 811

Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
 ||
 Db 812 TTTTGTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCAACAG 871

Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420
 |||||
 Db 872 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC 931

Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
 |||||
 Db 932 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGGCCACCATCAGCGTGATCATTGATGTGGAT 991

Qy 481 ATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
 ||
 Db 992 GTGAACATATCGGTTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG 1051

Qy 541 CTCTATTCTGTGGCCTACACTGATGTGCTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG 600
 |||||
 Db 1052 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTTATAGGACTGTGG 1111

Qy 601 ATCAGCGTCCCCTTTTGCATTGTGCATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
 |||||
 Db 1112 ATCAGTGTCCCCTTTTGGCCTGTGCATCCTGCAGTCACCGACATCGGATTCACAGCTGTG 1171

Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720
 |||||
 Db 1172 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG 1231

Qy 721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG 780
 |||||
 Db 1232 CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 1291

Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840
 ||
 Db 1292 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTCGGG 1351

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900
 |||||
 Db 1352 TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG 1411

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960

Db	1412	 AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	1471
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	1472	 CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1531
Qy	1021	TCTGCTGCTGTTATGTGCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCCATGTTTGCA	1080
Db	1532	 TCAGCTGCTGTGCATGTCCTCAGCTGACTCGTCCATCCTGTGCGGCGAGTTCTATGTTTGCT	1591
Qy	1081	CGGAACATCTACCAGCTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1592	 CGGAATATCTACCAGCTTTCTTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1651
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1652	 ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1711
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1712	 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG	1771
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1772	 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1831
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1832	 TTTGGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGACGCCCTTAATC	1891
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1892	 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1951
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1952	 ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	2011
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	2012	 CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	2071
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	2072	 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAAATGAAAATATCAAATTAAAT	2131
Qy	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	2132	 GAACCTTGCACTGTGAAACCTCGGCAGAGCCTAACCCTCAGTTCAACTTTCACCAATAAG	2191
Qy	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	2192	 GAGGCCCTTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAACTTACAA	2251
Qy	1741	TGA	1743

RESULT 4

AK034415

LOCUS AK034415 4306 bp mRNA linear HTC 18-SEP-2003

DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330188K24 product:solute carrier family 5 (choline transporter), member 7, full insert sequence.

ACCESSION AK034415

VERSION AK034415.1 GI:26329926

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 4306)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,

Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source

1. .4306
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9330188K24"
/db_xref="MGI:2398619"
/db_xref="taxon:10090"
/clone="9330188K24"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

CDS

394. .2136
/note="unnamed protein product; putative
solute carrier family 5 (choline transporter), member 7
(MGD|MGI:1927126, GB|NM_022025, evidence: BLASTN, 99%,
match=1743)"
/codon_start=1
/protein_id="BAC28702.1"
/db_xref="GI:26329927"
/translation="MSFHVEGLVAIILFYLLIFLVGIWAAWKTKNSGNPEERSEAIIV
GGRDIGLLVGGFTMTATWVG GYINGTAEAVYGP GCLAWAQAPIGYSLSLIIGGLFF
AKPMRSKGYVTMLDPFQQIYGKRMGGLLFI PALMGEMFWAAAI FSALGATISVIIDVD
VNISVIVSALIAILYTLVGGLYSVAYTDVVQLFCIFIGLWISVPFALSHPAVTDIGFT
AVHAKYQSPWLGTIESVEVYTWLDN FLLMLGGIPWQAYFQ RVLSSSSATYAQVLSFL
AAFGCLVMALPAICIGAIGASTDWNQTAYGYPDPKTKEEADMILPIVLQYLCPVYISF
FGLGAVSAAMSSADSSILSASSMFARNIYQLSFRQNASDKEI VWMRITVLVFGASA
TAMALLTKTVYGLWYLS SLDVYIIIFPQLLCVLFIKGTNTYGAVAGYIFGLFLRITGG
EPYLYLQPLIFYPGYSDKNGIYNQRFPFKTL SMVTSFFTNCVSYLAKYLFESGTL P
PKLDVFD AVVARHSEENMDKTI LVRNENIKL NELAPVKPRQSLT LSSFTNKEALLDV

DSSPEGSGETEDNLQ"

ORIGIN

Query Match 78.9%; Score 1375; DB 11; Length 4306;
 Best Local Similarity 86.8%; Pred. No. 0;
 Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||| || ||||| ||||| ||| | ||||| || || || |||
Db     394 ATGTCCTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 453

Qy      61 GTTGGAAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db     454 GTTGGAAATATGGGCTGCATGGAAAACCAAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 513

Qy     121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
Db     514 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC 573

Qy     181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||
Db     574 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGATGGGCCAGGTTGT 633

Qy     241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db     634 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 693

Qy     301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db     694 TTTTGTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCACAG 753

Qy     361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420
      ||||| ||||| ||||| || || || || || || || || || |||||
Db     754 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT 813

Qy     421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      ||||| ||||| ||||| || || ||||| ||||| ||||| ||||| |||
Db     814 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT 873

Qy     481 ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      || |||| || ||||| ||||| ||||| ||||| || || || || || |||
Db     874 GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG 933

Qy     541 CTCTATTCTGTGGCCTACACTGATGTCGTTGAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600
      ||||| ||||| || ||||| || ||||| || ||||| ||||| |||||
Db     934 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 993

Qy     601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     994 ATCAGTGTCCCTTTTGCCCTGTCACATCCTGCAGTCACCGACATCGGATTACAGCTGTG 1053

Qy     661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720
      ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| |||
Db    1054 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG 1113

Qy     721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAAGAG 780
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	1114	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	1173
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTTCGGG	840
Db	1174	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG	1233
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1234	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	1293
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1294	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	1353
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTGGTCTTGGTGCA	1020
Db	1354	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTGGGCTTGGTGCTGTT	1413
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCCATGTTTGCA	1080
Db	1414	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGGCGAGTTCTATGTTTGCT	1473
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1474	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1533
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1534	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1593
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1594	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTTTACATCATCATCTTCCCACAG	1653
Qy	1261	CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1654	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1713
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1714	TTTGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGACGCCCTTAATC	1773
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1774	TTCTACCCTGGTTATTACTCTGACAAGAAATGGTATATACAATCAGAGGTTCCCATTTAAA	1833
Qy	1441	ACACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1834	ACTCTCTCCATGGTTACCTCATCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1893
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1894	CTATTTGAAAGTGGAACCTTGCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	1953
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1954	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT	2013

Qy 1621 GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
 ||||| ||||| || || ||||| | ||||| ||||| |||||
 Db 2014 GAACTTGCACCTGTGAAACCTCGGCAGAGCCTAACCTCAGTTCAACTTTCACCAATAAG 2073

Qy 1681 GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2074 GAGGCCCTTCCTTGATGTTGATTCCAGTCCGGAGGGTCTGGGACTGAAGATAACTTACAA 2133

Qy 1741 TGA 1743
 |||
 Db 2134 TGA 2136

RESULT 5

AY413300

LOCUS AY413300 1743 bp DNA linear GSS 12-DEC-2003

DEFINITION Mus musculus HCM4844 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY413300

VERSION AY413300.1 GI:39769262

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..1743
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 gene <1..>1743
 /locus_tag="HCM4844"

ORIGIN

Query Match 66.3%; Score 1156.2; DB 29; Length 1743;
 Best Local Similarity 73.1%; Pred. No. 3.8e-304;
 Matches 1275; Conservative 0; Mismatches 468; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
 ||| ||||| || ||||| ||||| ||| | ||||| || || || |||
 Db 1 ATGTCTTTCCACGTAGAAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 60

Qy 61 GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 61 GTTGGAATATGGGCTGCATGGAAAACCAAAACAGCGGCAACCCAGAAGAGCACAGTGAA 120

Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
 ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 121 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGNN 180

Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
 Db 181 NNN 240

Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
 Db 241 NNN 300

Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
 Db 301 NNN 360

Qy 361 ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420
 Db 361 NNN 420

Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 421 NNGGGCCACCATCAGCGTGATCATCGATGTGGAT 480

Qy 481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
 || |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 481 GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCCTAGTGGGTGGG 540

Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTGAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600
 ||||| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| |||||
 Db 541 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 600

Qy 601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGAGACATCGGGTTCACTGCTGTG 660
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 601 ATCAGTGTCCCCTTTTGCCCTGTCACATCCTGCAGTCACCGACATCGGATTACAGCTGTG 660

Qy 661 CATGCCAAATACCAAAAAGCCGTGGCTGGGAACTGTGACTCATCTGAAGTCTACTCTTGG 720
 ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 661 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG 720

Qy 721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 721 CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 780

Qy 781 GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTTCGGG 840
 || ||||| ||||| ||||| ||||| || ||||| ||||| ||||| ||||| |||||
 Db 781 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG 840

Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTTACATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTATTTCCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCTCCAAAATTAGATGTATTTGATGCTGTTGTGCAAGG	1560
Qy	1561	CACAGTGAAGAAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATCTAGTCAGAAATGAAAATATCAAATTAAAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGACCTGTGAAACCTCGGCAGAGCCTAACCTCAGTTCAACTTTCACCAATAAG	1680
Qy	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740

```

      ||||| |||||||||||||||||||| || |||||||||||||||||||| |||||
Db      1681 GAGGCCCTCCTTGATGTTGATTCCAGTCCGGAGGGGTCTGGGACTGAAGATAACTTACAA 1740

Qy      1741 TGA 1743
      |||
Db      1741 TGA 1743

```

RESULT 6

AG157499

LOCUS AG157499 672 bp DNA linear GSS 09-JAN-2002

DEFINITION Pan troglodytes DNA, clone: RP43-022H02.T7, genomic survey sequence.

ACCESSION AG157499

VERSION AG157499.1 GI:16687177

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE BAC end sequences of Library RPCI-43

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 672)

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT Clones are derived from the chimpanzee BAC library RPCI-43 This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: T7

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

FEATURES

source

Location/Qualifiers

1. .672

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="RP43-022H02.T7"

/sex="male"

/cell_type="lymphocytes"

/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 29.8%; Score 518.8; DB 29; Length 672;

Best Local Similarity 97.3%; Pred. No. 3.5e-130;

Matches 549; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

Accession	Position	Sequence	Length
Db	207	ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	266
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	267	GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	326
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	327	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCAACAG	386
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGG-AGAAATGTT	419
Db	387	ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGNAGAGATGTT	446
Qy	420	CTGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCA-TCAGCGTGATCA-TCGATGTG	477
Db	447	CTGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATTGAGCGTGATCATTGGATGTG	506
Qy	478	GATATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGA	537
Db	507	GATGTGAACATATCGGTTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGT	566
Qy	538	GGGCTCTATTCTGTGGCCTACACT-GATGTGCTTCAGCTCTTTTGCATTTTGTAGGGCT	596
Db	567	GGGCTCTACTCTGTGGCATATACTGGATGTTGTCCAGCTATTCTGCATTTTATAGGACT	626
Qy	597	GT-GGATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTG	655
Db	627	GTGGGATCAGTGTCCCTTTTGCCTGTGCACATCCTGCAGTCACCGACATCGGATTTCACAG	686
Qy	656	CTGTGCATGCCAA	668
Db	687	CTGTGCATGCTAA	699

Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
 Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
 Kawaji,H., Kawasaki,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
 Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
 Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
 Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G.,
 Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
 Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
 Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
 Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
 Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
 Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
 Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 22354683

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
 Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
 Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
 Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
 Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
 Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
 Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
 Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source Location/Qualifiers
1. .669
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="B730003H24"
/tissue_type="medulla oblongata"
/dev_stage="6 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 6 days neonate medulla oblongata"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 26.6%; Score 462.8; DB 13; Length 669;
Best Local Similarity 86.7%; Pred. No. 6.8e-115;
Matches 509; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||| || ||||| ||||| ||| | ||||| || || || |||
Db      81 ATGTCTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTCTG 140

Qy      61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db     141 GTTGAATATGGGCTGCATGGAAAACCAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 200

Qy     121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||
Db     201 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGGTGGTGGTGGTTTTACCATGACAGCC 260

Qy     181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||
Db     261 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 320

Qy     241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db     321 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 380

Qy     301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
```



```

/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1PIG"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

```

ORIGIN

```

Query Match          23.2%;  Score 404;  DB 10;  Length 516;
Best Local Similarity 86.4%;  Pred. No. 7.2e-99;
Matches 446;  Conservative 0;  Mismatches 70;  Indels 0;  Gaps 0;

Qy      1067 GTTCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAG 1126
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db        1 GTTCTATGTTTGTAGAAACATCTAGCAGCTCTCATTGACACAAAACGCTTCCGACAGGG 60

Qy      1127 AAATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGG 1186
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db        61 AGATCGTCTGGGTCATGCGGATCACAGTATTCGTGTTTGGTGCGTCTGCAACAGCCATGG 120

Qy      1187 CCTTGCTGACGAAAACGTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCG 1246
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       121 CCTTGCTGACCAAGACCGTGTATGGGCTCTGGTACCTCAGTCCGACCTCGTCTACATCA 180

Qy      1247 TTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCG 1306
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       181 TTATCTTCCCGCAGCTGCTCTGTGTGCTCTTCATCAAGGGGACCAACACGTACGGGGCCG 240

Qy      1307 TGGCAGGTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATC 1366
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       241 TGGCAGGGTACATTGCTGGCCTTTTCCTGAGGGTAACCGGTGGAGAGCCATACCTGAACC 300

Qy      1367 TTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGA 1426
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       301 TGCAGCCCTTGATCTTTTACCCTGGTTATTACGTTGAAAAAAATGGTATATATAATCAGA 360

Qy      1427 AATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCT 1486
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       361 GATTCCCATTTAAAACCTTGCCATGCTCACCTCCTTCTTATCCAACATTTGCATCTCTT 420

Qy      1487 ATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATG 1546
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       421 ATCTAGCCAAATATCTATTTGAAAGTGGAACCTTGCCACCAAAATTAGATATGTTTGATG 480

Qy      1547 CTGTTGTTGCAAGACACAGTGAAGAAAACATGGATA 1582
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       481 CTGTTGTTGCAAGACACAGTGAAGAAAACATGGATA 516

```

RESULT 10

BB626260

LOCUS BB626260 650 bp mRNA linear EST 26-OCT-2001

DEFINITION BB626260 RIKEN full-length enriched, adult male diencephalon Mus musculus cDNA clone 9330170D24 5', mRNA sequence.

ACCESSION BB626260

VERSION BB626260.1 GI:16464298
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 650)
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
 Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
 Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
 Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
 Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
 TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 JOURNAL Unpublished (2001)
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. . 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
 Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.
 FEATURES Location/Qualifiers
 source 1. .650
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="9330170D24"

Qy 421 TGGGCTGCAGCAATTT 436
 |||||
 Db 620 TGGGCTGCAGCCATTT 635

RESULT 11

AW668962

LOCUS AW668962 541 bp mRNA linear EST 25-APR-2001

DEFINITION 111664 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION AW668962

VERSION AW668962.1 GI:7525476

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 541)

AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
 Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
 Chitko-McKown,C.G., Perteau,G., Holt,I., Karamycheva,S., Liang,F.,
 Quackenbush,J. and Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

PUBMED 11282978

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGACG

Plate: 95 row: L column: 20

Seq primer: ATTTAGGTGACACTATAG.

FEATURES Location/Qualifiers

source

1. .541

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 1BOV"

/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;

Library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."

ORIGIN

Query Match

17.9%; Score 312.8; DB 10; Length 541;

Best Local Similarity 85.0%; Pred. No. 5.9e-74;
Matches 350; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```

Qy      889 TCAACAGACTGGAACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCA 948
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      130 TCAACAGCCTGGAACCAGACTGCATACGGGCTCTTGCTCCCAGGGAGAAACAGGAGGCA 189

Qy      949 GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT 1008
        ||||| || || ||||| || ||||| ||||| ||||| ||||| |||||
Db      190 GACATGATCTTGCCGATTGTCTCAAGTATCTCTGCCCCGTGTACATTTCTTACTTTGGT 249

Qy     1009 CTTGGTGCACTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGT 1068
        ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      250 CTTGGAGCCGTTTCTGCTGCTGTCATGTCCTCAGCAGATTCTTCCATCTTGTCAGCAAGT 309

Qy     1069 TCCATGTTTGACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAA 1128
        || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      310 TCGATGTTTGCTCGCAACATCTACCAGCTTTCATTTCAGACAAAATGCTTCTGACAAGGAG 369

Qy     1129 ATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCC 1188
        || || ||||| ||||| ||||| || ||||| ||||| ||||| || |||||
Db      370 ATAGTCTGGGTCATGCGCATCACGGTATTTGTGTTTGGAGCTTCTGCGATGACCATGGCC 429

Qy     1189 TTGCTGACGAAAACCTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTT 1248
        ||||| ||||| || ||||| ||||| ||||| ||||| ||||| || |||||
Db      430 TTGCTAACGAAGACGGTGTATGGGCTCTGGTACCTCAGTCTGACCTGGTCTACATCATC 489

Qy     1249 ATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATG 1300
        ||||| || ||||| || || ||||| || ||||| ||||| |||||
Db      490 ATCTTCCCGCAGTTGCTCTGCGTGCTCTTCATCAAGGGTACCAACACGTATG 541

```

RESULT 12

BY729567

LOCUS BY729567 675 bp mRNA linear EST 17-DEC-2002

DEFINITION BY729567 RIKEN full-length enriched, 12 days embryo spinal cord Mus musculus cDNA clone C530033E06 5', mRNA sequence.

ACCESSION BY729567

VERSION BY729567.1 GI:27142694

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 675)

AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusica,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustinich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,

Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
 Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G.,
 Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
 Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
 Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
 Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
 Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
 Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
 Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 22354683

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
 Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
 Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
 Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
 Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
 Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
 Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
 Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

FEATURES Location/Qualifiers

source 1. .675

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="C530033E06"
/tissue_type="spinal cord"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo
spinal cord"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAAGGATCCAAGAGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
was cleaved with XhoI and BamHI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

```

ORIGIN

```

Query Match          16.6%;  Score 290;  DB 13;  Length 675;
Best Local Similarity 86.5%;  Pred. No. 1.1e-67;
Matches 320;  Conservative 0;  Mismatches 50;  Indels 0;  Gaps 0;

```

```

Qy      889 TCAACAGACTGGAACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCA 948
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      306 TCCACAGACTGGAACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCA 365

Qy      949 GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT 1008
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      366 GACATGATTCTCCCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGG 425

Qy     1009 CTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGT 1068
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      426 CTTGGTGTGTTTCAGCTGCTGTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGT 485

Qy     1069 TCCATGTTTGCACGGAACATCTACCAGCTTTCCCTTCAGACAAAATGCTTCGGACAAAGAA 1128
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      486 TCTATGTTTGTCTCGGAATATCTACCAGCTTTCCCTTCAGACAAAATGCATCAGACAAGGAA 545

Qy     1129 ATCGTTTGGGTTATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCC 1188
      || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      546 ATTGTGTGGGTCATGAGGATCACTGTGCTTGTGTTGCGAGCATCTGCAACAGCCATGGCT 605

Qy     1189 TTGCTGACGAAAACGTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTT 1248
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      606 TTGCTGACGAAGACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTTTACATCATC 665

Qy     1249 ATCTTCCCCC 1258
      ||||| |||||
Db      666 ATCTTCCCAC 675

```

RESULT 13
 BE723927
 LOCUS BE723927 524 bp mRNA linear EST 25-APR-2001
 DEFINITION 198406 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BE723927
 VERSION BE723927.1 GI:10125223
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 524)
 AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
 Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
 Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
 Quackenbush,J. and Keele,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013
 PUBMED 11282978
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAAACAGCTATGACCAT
 BACKWARD: GTTTTCCCAGTCACGACG
 Plate: 106 row: L column: 14
 Seq primer: ATTTAGGTGACACTATAG.
 FEATURES Location/Qualifiers
 source 1. .524
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 4BOV"
 /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
 Library made from pooled tissue from day 20 and day 40
 embryos."
 ORIGIN
 Query Match 15.7%; Score 274.2; DB 10; Length 524;
 Best Local Similarity 84.7%; Pred. No. 2e-63;
 Matches 331; Conservative 0; Mismatches 58; Indels 2; Gaps 2;
 Qy 889 TCAACAGACTGGAACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCA 948

```

      |||||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      130 TCAACATCCTGGAACCAGACTGCATACGGGCCTCTTGCTCCCAGGGAGAAACAGGAGGCA 189

Qy      949 GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT 1008
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      190 GACATGATCTTGCCGATTGTCCCTCAAGTATCTCTGCCCCGTGTACATTTCTTACTTTGGT 249

Qy      1009 CTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTC-TTCCATCTTGTGCAAG 1067
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      250 CTTGGAGCCGTTTCTGCTGCTGTATGTCCTCAGCAGATTCCTTCCATCTTGTGCAAG 309

Qy      1068 TTCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGA 1127
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      310 TTCGATGTTTGTCTCGCAACATCTACCAGCTTTCATTTCAGACAAAATGCTTCTGACAAGGA 369

Qy      1128 AATCGTTTGGGTTATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGC 1187
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      370 GATAGTCTGGGTCATGCGCATCACGGTATTTGTGTTGGAGCTTCTGCGATGACCATGGC 429

Qy      1188 CTTGCTGACGAAAACGTGTGTATGGGCTC-TGGTACCTCAGTTCTGACCTTGTGTTACATCG 1246
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      430 CTTGCTAACGAAGACGGTGTATGGGCTCTTGGTACCTCAGCTCTGACCTGGTCTACATCA 489

Qy      1247 TTATCTTCCCCCAGCTGCTTTGTGTACTCTT 1277
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      490 TCATCTTCCCGCAGTTGCTCTGCGTGCTCTT 520

```

RESULT 14

AL669749

LOCUS AL669749 800 bp mRNA linear EST 14-JAN-2002

DEFINITION AL669749 directional larval cDNA library *Ciona intestinalis* cDNA clone 052ZB03 5', mRNA sequence.

ACCESSION AL669749

VERSION AL669749.1 GI:18143007

KEYWORDS EST.

SOURCE *Ciona intestinalis*

ORGANISM *Ciona intestinalis*

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; *Ciona*.

REFERENCE 1 (bases 1 to 800)

AUTHORS Genoscope.

TITLE *Ciona intestinalis* directional larval cDNA library

JOURNAL Unpublished (2002)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

IMPORTANT: this sequence may contain errors. The *Ciona intestinalis* library from which the clone was isolated may be contaminated with cDNAs from bacteria or other Eukarya.

Directional larval cDNA library originate from Dr.M.Branno, Stazione A.Dohrn, Naples, Italy, and was prepared in pBluescript2SK+.

FEATURES Location/Qualifiers

source 1..800

/organism="Ciona intestinalis"

/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="052ZB03"
/clone_lib="directional larval cDNA library"
/note="Vector: pBluescript2SK+"

ORIGIN

Query Match 15.1%; Score 263.2; DB 9; Length 800;
Best Local Similarity 62.7%; Pred. No. 2.3e-60;
Matches 504; Conservative 0; Mismatches 286; Indels 14; Gaps 6;

Qy 250 TGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGC- 308
||| | || ||||| ||||| || | || | || | ||||| |||||
Db 2 TGGACGCAAGCACCCATTGGATACGCTTGCGCGTTAATACTTGGCGGCTTATTCTTTGCG 61

Qy 309 AAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGG 368
|| ||||| ||||| ||||| ||||| || || | ||||| | || |||
Db 62 AAGTAAAATGCGAAGTGAGGGATATGTGACGATGTTGGATCCACTGCAGCGCAACT-TGG 120

Qy 369 AAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTCTGGGCTGC 428
| ||||| | || | || ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TCGAGTAATGGGAGCGTTTCTTTATATACCTGCACTTGCTGGAGAATTATTCTGGTCTGC 180

Qy 429 AGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCACAT 488
||| || || | || ||||| | ||| ||| || ||||| ||| | |||||
Db 181 AGCTATATTGGCCGCGTTGGGCGGTACCTTCA-TGTTATCATTGATCTTCATATAACTGC 239

Qy 489 TTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTC 548
|||| || | ||||| ||||| | ||||| ||| || || || || || ||
Db 240 AGCTGTAATAGTATCTGCATGCATTGCTGTTGTATACACCATGGCCGGTGGTCTTTACTC 299

Qy 549 TGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGGATCAGCGT 608
|| || || || ||||| ||||| | ||||| || | || ||||| | ||| |
Db 300 GGTGCTTATACAGATGTAGTTCAGTTGATTGTCATATTCATTGGACTGTGGTTGAGCAT 359

Qy 609 CCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTGCATGCCAA 668
|| || || || || ||||| || ||||| || ||||| || ||||| || |||
Db 360 TCCATTGCGGTTCACTCATCCTGCTGTATCAGACATCGCCACTACAGCTTACCACTCACC 419

Qy 669 ATACCAAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGGCTTGATAG 728
|| || ||||| || ||| || || || || || || || || || || ||
Db 420 TAAC-----TGGCTTGGTACTTGGGATATTCGACCACTGGTCTATGGATCGACTC 470

Qy 729 TTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGGGTTCTCTC 788
| |||| | || | ||||| || ||||| ||||| || ||||| || |||
Db 471 TGCTCTGCTACTGTTATTTGGTGGAAATACCGTGGCAAGTTTACTTTCAAAGAGTTTATC 530

Qy 789 TTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGGTGCCTGGT 848
| | | |||| ||| || || || || || || || || || || || ||
Db 531 GNCNAAAAGNNCANGAAGCGCTCAGAAGCTTTCATTGCTGCGTTGCGATGTTTGTGTT 590

Qy 849 GATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAACCAGAC 908
||| | || || || || || || || || || || || || || || || ||
Db 591 CATGTCAATACCTTCGATATTNATCGGTGCAATTGCTGCATCTACAGATTGGGACGCAAC 650

Qy 909 TGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTACCAATTGT 968
| || || || || || || || || || || || || || || || || || || || ||


```

Db          651 ATCGTACGGCCTCCCAAGTCCAGTTGANAAAGGCGACCAAGCCAATATTCTACCCATTGT 710
Qy          969 TCTGCAGTATCTC-TGCCCTGTGTATATTTCTTTCTTT-GGTCTTGGTGCAGTTTCTGCT 1026
              || || || ||| ||||| | | || ||||| || ||||| || ||||| |||
Db          711 GCTTCAATACCTCACCCCTGTAGCTGTATCATTCCTTTGGGGCTTGGCGCTGTTTCTGCT 770
Qy          1027 GCTGTTATGTCATCAGCAGATTCT 1050
              ||||| ||||| || || |||
Db          771 GCTGTNATGTCATCTGCCGACTCT 794

```

RESULT 15

BW274870

LOCUS BW274870 549 bp mRNA linear EST 11-NOV-2002

DEFINITION BW274870 Nori Satoh unpublished cDNA library, gastrula and neurula
Ciona intestinalis cDNA clone cign070c12 5', mRNA sequence.

ACCESSION BW274870

VERSION BW274870.1 GI:24855481

KEYWORDS EST.

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 549)

AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.

TITLE Expressed genes in Ciona intestinalis (2002c)

JOURNAL Unpublished (2002)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source

Location/Qualifiers

1. .549

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="cign070c12"

/tissue_type="whole body"

/dev_stage="gastrula and neurula"

/clone_lib="Nori Satoh unpublished cDNA library, gastrula
and neurula"

ORIGIN

Query Match 13.0%; Score 225.8; DB 13; Length 549;

Best Local Similarity 64.4%; Pred. No. 3.4e-50;

Matches 354; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

```

Qy          11 ATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTGGTTGGAATAT 70
              ||| | || | | ||| ||||| ||||| || | || | |||
Db          2 ATGTTCTGTTTAGTGNCATTATCGTCTTCTACGTTGCTATTCTAGCGATCGGTATTT 61
Qy          71 GGGCTGCCTGGAGAACCAAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATCATAG 130
              || || ||||| | | || | | || || | ||| ||||| | ||||| |
Db          62 ATGCAGCATGGAGGAAAAAGAAGAACCGGAAGAGGAAACGAG---AGCGAGACAATCATGG 118

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2004, 20:07:21 ; Search time 10873 Seconds
(without alignments)
6948.120 Million cell updates/sec

Title: US-10-069-541-5
Perfect score: 1743
Sequence: 1 atggcctttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

28: em_un:*
 29: em_vi:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_pln:*
 35: em_htg_rod:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	8						Description
	No.	Score	Query Match	Length	DB	ID	
	1	1743	100.0	1743	6	E49871	E49871 High-affini
	2	1743	100.0	1743	6	BD012719	BD012719 High-affi
	3	1743	100.0	1743	9	AF276871	AF276871 Homo sapi
	4	1743	100.0	1813	9	HSA401466	AJ401466 Homo sapi
	5	1743	100.0	5158	9	AB043997	AB043997 Homo sapi
	6	1738.2	99.7	1743	6	AR268949	AR268949 Sequence
	7	1394.2	80.0	1743	6	E49870	E49870 High-affini
	8	1394.2	80.0	1743	6	BD012718	BD012718 High-affi
	9	1394.2	80.0	4904	10	AB030947	AB030947 Rattus no
	10	1375	78.9	1743	10	AF276872	AF276872 Mus muscu
	11	1373.4	78.8	1743	6	E49872	E49872 High-affini
	12	1373.4	78.8	1743	6	BD012720	BD012720 High-affi
	13	1373.4	78.8	4938	6	AX080443	AX080443 Sequence
	14	1367	78.4	1743	10	MMU401467	AJ401467 Mus muscu
	15	867	49.7	2528	5	TMA420808	AJ420808 Torpedo m
	16	730	41.9	1132	5	GGA511267	AJ511267 Gallus ga
	17	630.8	36.2	2239	9	HSA308384	AJ308384 Homo sapi
	18	630.8	36.2	190043	9	AC009963	AC009963 Homo sapi
	19	502.8	28.8	232792	2	AC106657	AC106657 Rattus no
c	20	501.2	28.8	155131	2	AC102873	AC102873 Mus muscu
	21	431.4	24.8	3326	3	AY011119	AY011119 Limulus p
	22	405.8	23.3	3255	3	AY047521	AY047521 Drosophil
	23	363.8	20.9	1731	6	E49869	E49869 High-affini
	24	363.8	20.9	1731	6	BD012717	BD012717 High-affi
	25	363.8	20.9	1985	3	AB030946	AB030946 Caenorhab
	26	279.6	16.0	386	6	AX080449	AX080449 Sequence
	27	242.6	13.9	1461	6	AX432086	AX432086 Sequence
	28	226	13.0	1657	9	HSA308383	AJ308383 Homo sapi
	29	179.6	10.3	1178	9	HSA308378	AJ308378 Homo sapi
	30	179.6	10.3	186989	3	AC007812	AC007812 Drosophil
	31	179.6	10.3	189117	3	AC009395	AC009395 Drosophil
	32	179.6	10.3	255620	3	AE003723	AE003723 Drosophil
c	33	167.6	9.6	140156	2	AC017381	AC017381 Drosophil

	34	163	9.4	2326	9	HSA308379	AJ308379 Homo sapi
	35	155	8.9	1467	9	HSA308382	AJ308382 Homo sapi
c	36	151.8	8.7	40893	3	CBRG45E19	AC084631 Caenorhab
	37	150.2	8.6	736	9	HSA308381	AJ308381 Homo sapi
	38	150	8.6	1308	9	HSA308380	AJ308380 Homo sapi
c	39	141.6	8.1	39908	3	CEC48D1	Z81049 Caenorhabdi
c	40	141.6	8.1	330724	2	CEY67H2	AL022475 Caenorhab
c	41	132	7.6	152021	2	AC010923	AC010923 Drosophil
	42	98	5.6	616	11	G84799	G84799 S208P6036FB
c	43	51.8	3.0	10732	6	E32986	E32986 Gene encodi
c	44	49	2.8	2000	6	AX655393	AX655393 Sequence
	45	48.2	2.8	2781	1	BSU92466	U92466 Bacillus su

ALIGNMENTS

RESULT 1

E49871

LOCUS E49871 1743 bp DNA linear PAT 27-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION E49871

VERSION E49871.1 GI:22554902

KEYWORDS JP 2001136976-A/3.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Haga,T. and Okuda,T.

TITLE High-affinity choline transporter

JOURNAL Patent: JP 2001136976-A 3 22-MAY-2001;

SCIENCE & TECH AGENCY

COMMENT OS Homo sapiens (human)

PN JP 2001136976-A/3

PD 22-MAY-2001

PF 27-DEC-1999 JP 1999368991

PI TATSUYA HAGA,TAKASHI OKUDA

PC C12N15/09,A01K67/027,A61K38/00,C07K14/47,C07K16/18,C07K19/00,

PC C12N5/10,

PC C12P21/02,C12P21/08,C12Q1/00,C12N15/00,A61K37/02,C12N5/00 CC

FH Key Location/Qualifiers

FT CDS (1)..(1743).

FEATURES Location/Qualifiers

source 1..1743

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1743; DB 6; Length 1743;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

|||||

Db 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Qy	181	ACCTGGGTTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900

Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743

Db 1741 TGA 1743

RESULT 2

BD012719

LOCUS BD012719 1743 bp DNA linear PAT 02-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION BD012719

VERSION BD012719.1 GI:22092908

KEYWORDS WO 0116315-A/3.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Haga,T. and Okuda,T.

TITLE High-affinity choline transporter

JOURNAL Patent: WO 0116315-A 3 08-MAR-2001;

JAPAN SCIENCE AND TECHNOLOGY CORP,TATSUYA HAGA,TAKASHI OKUDA

COMMENT OS Homo sapiens (human)

PN WO 0116315-A/3

PD 08-MAR-2001

PF 18-AUG-2000 WO 2000JP005545

PR 27-AUG-1999 JP 99P 240642,27-DEC-1999 JP 99P 368991 PI

TATSUYA HAGA,TAKASHI OKUDA

PC C12N15/12,C07K14/47,C12Q1/68,C07K19/00,C07K16/18,C12N5/10, PC
A61K38/17,

PC A61K45/00,A61P25/28,G01N33/53,A01K67/027

CC

FH Key Location/Qualifiers

FT CDS (1)..(1743).

FEATURES Location/Qualifiers

source 1..1743

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1743; DB 6; Length 1743;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
          |||
Db      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy     61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
          |||
Db     61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
          |||
Db    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180

Qy    181 ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
          |||
```


Db	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTCTTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTCTTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080

Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 3

AF276871

LOCUS AF276871 1743 bp mRNA linear PRI 27-NOV-2000

DEFINITION Homo sapiens high affinity choline transporter (SLC5A7) mRNA, complete cds.

ACCESSION AF276871

VERSION AF276871.1 GI:10998441

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Apparsundaram, S., Ferguson, S.M., George, A.L. Jr. and Blakely, R.D.

TITLE Molecular cloning of a human, hemicholinium-3-sensitive choline transporter

JOURNAL Biochem. Biophys. Res. Commun. 276 (3), 862-867 (2000)

MEDLINE 20483599

PUBMED 11027560

REFERENCE 2 (bases 1 to 1743)

AUTHORS Apparsundaram, S., Ferguson, S.M. and Blakely, R.D.

TITLE Direct Submission

JOURNAL Submitted (09-JUN-2000) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at Pierce, Nashville, TN 37232-6420, USA

FEATURES Location/Qualifiers

source 1. .1743
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q12"

gene 1. .1743
/gene="SLC5A7"

CDS 1. .1743
/gene="SLC5A7"
/note="hCHT; solute carrier family 5 member 7"
/codon_start=1
/product="high affinity choline transporter"
/protein_id="AAG25940.1"
/db_xref="GI:10998442"
/translation="MAFHVEGLIAIIVFYLLILLVGIWAAWRTKNSGSAEERSEAIIV
GGRDIGLLVGGFTMTATWVG GGYINGTAEAVYVPGYGLAWAQAPIGYSLSLILGGLFF
AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAI FSALGATISVIIDVD
MHISVIISALIATLYTLVGGLYSVAYTDVVQLFCIFVGLWISVPFALSHPAVADIGFT
AVHAKYQKPWLGTVDSSSEVYSWLDSFLLLMLGGIPWQAYFQRLSSSSATYAQVLSFL
AAFGCLVMAIPAILIGAIGASTDWNQTAYGLPDPKTEEADMILPIVLQYLCVPYISF
FGLGAVSAVMSSADSSILSASSMFARNIYQLSFRQNASDKETVWVMRITVFVFGASA
TAMALLTKTVYGLWYLSSDLVYIVIFPQLLCVLFVKGTNTYGAVAGYVSGLFLRITGG
EPYLYLQPLIFYPGYPPDDNGIYNQKFPFKTLAMVTSFLTNICISYLA KYLFESGTLF
PKLDVFDVAVVARHSEENMDKTIILVKNENIKLDELALVKPRQSMTLSSTFTNKEAFLDV
DSSPEGSGT EDNLQ"

ORIGIN

Query Match 100.0%; Score 1743; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
|||||

Db 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
|||||

Db 61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
 |||

Db 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180

Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
 |||

Db 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240

Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
 |||

Db 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300

Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
 |||

Db 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360

Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420
 |||

Db 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420

Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
 |||

Db 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480

Qy 481 ATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
 |||

Db 481 ATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540

Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600
 |||

Db 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600

Qy 601 ATCAGCGTCCCCTTTGCATTGTTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
 |||

Db 601 ATCAGCGTCCCCTTTGCATTGTTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660

Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG 720
 |||

Db 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG 720

Qy 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
 |||

Db 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780

Qy 781 GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG 840
 |||

Db 781 GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG 840

Qy 841 TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900
 |||

Db 841 TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960
 |||

Db 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960

Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGT	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 4

HSA401466

LOCUS HSA401466 1813 bp mRNA linear PRI 16-AUG-2000

DEFINITION Homo sapiens mRNA for high affinity choline transporter (CHT1 gene).

ACCESSION AJ401466

VERSION AJ401466.1 GI:9843753

KEYWORDS ChT1 gene; high affinity choline transporter.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Wieland,A., Bonisch,H. and Bruss,M.

TITLE Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1813)

AUTHORS Bruess,M.

TITLE Direct Submission

JOURNAL Submitted (14-AUG-2000) Bruess M., University of Bonn, Pharmacology and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY

FEATURES Location/Qualifiers

source

1. .1813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q11-13"
/tissue_type="hypothalamus"

gene

1. .1813
/gene="CHT1"

CDS

19. .1761
/gene="CHT1"
/function="sodium- and chloride-dependent reuptake of choline"
/codon_start=1
/evidence=experimental
/product="high affinity choline transporter"
/protein_id="CAC03717.1"
/db_xref="GI:9843754"
/db_xref="GOA:Q9GZV3"
/db_xref="SPTREMBL:Q9GZV3"
/translation="MAFHVEGLIAIIVFYLLILLVGIWAAWRTKNSGSAEERSEAIIV
GGRDIGLLVGGFTMTATWVG GGYINGTAEAVIVPGYGLAWAQAPIGYSLSLILGGLFF
AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAFSALGATISVIIDVD
MHISVIIISALIATLYTLVGGLYSVAYTDVVQLFCIFVGLWISVPFALSHPAVADIGFT
AVHAKYQKPWLGTVDSSSEVYSWLDSEFLMLLGGIPWQAYFQRLSSSSATYAQVLSFL
AAFGCLVMAIPAILIGAIGASTDWNQTAYGLPDPKTTTEADMILPIVLQYLCPVYISF
FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVMRITVFVFGASA
TAMALLTKTVYGLWYLSDDLVIYIVIFPQLLCVLFVKGTNTYGAAGYVSGFLRITGG
EPYLYLQPLIFYPGYYPDDNGIYNQKFPFKTLAMVTSFLTNICISYLA KYLFESGTLF
PKLDVFDVAVVARHSEENMDKTI L VKNENIKLDELALVKPRQSM T L S S T F T N K E A F L D V
DSSPEGSGTEDNLQ"

ORIGIN

Query Match 100.0%; Score 1743; DB 9; Length 1813;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	19	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	78
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	79	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	138
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Db	139	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	198
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	199	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	258
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	259	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	318
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	319	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	378
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Db	379	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	438
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	439	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	498
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	499	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	558
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	559	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	618
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	619	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	678
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	679	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	738
Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	739	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	798
Qy	781	GTTCTCTCTTCTCCTCAGCCACCTATGCTCAAGTGTCTCCTCCTGGCAGCTTTCGGG	840

Db	799	 GTTCTCTCTTCTTCCCTCAGCCACCTATGCTCAAGTGTCTCCTTCCTGGCAGCTTTCGGG	858
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	859	 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	918
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	919	 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	978
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	979	 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1038
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCTAGCAAGTTCCATGTTTGCA	1080
Db	1039	 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCTAGCAAGTTCCATGTTTGCA	1098
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1099	 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1158
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1159	 ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1218
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1219	 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1278
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1279	 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1338
Qy	1321	TCTGGCCTCTTCCTGAGAAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1339	 TCTGGCCTCTTCCTGAGAAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1398
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1399	 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1458
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1459	 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1518
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1519	 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1578
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1579	 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1638
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAAA	1680

Db 1639 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1698

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 |||

Db 1699 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1758

Qy 1741 TGA 1743
 |||

Db 1759 TGA 1761

RESULT 5
 AB043997

LOCUS AB043997 5158 bp mRNA linear PRI 19-NOV-2000

DEFINITION Homo sapiens mRNA for high-affinity choline transporter CHT1,
 complete cds.

ACCESSION AB043997

VERSION AB043997.1 GI:11231080

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Okuda,T. and Haga,T.

TITLE Functional characterization of the human high-affinity choline
 transporter

JOURNAL FEBS Lett. 484 (2), 92-97 (2000)

MEDLINE 20521663

PUBMED 11068039

REFERENCE 2 (bases 1 to 5158)

AUTHORS Okuda,T.

TITLE Direct Submission

JOURNAL Submitted (30-MAY-2000) Takashi Okuda, University of Tokyo, Faculty
 of Medicine, Department of Neurochemistry; 7-3-1 Hongo, Bunkyo-ku,
 Tokyo 1130033, Japan (E-mail:okuda@m.u-tokyo.ac.jp,
 URL:http://park.ecc.u-tokyo.ac.jp/neurochemistry,
 Tel:81-3-5841-3560, Fax:81-3-6814-8154)

FEATURES Location/Qualifiers

source

1. .5158

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="spinal cord"

CDS

277. .2019

/codon_start=1

/product="high-affinity choline transporter CHT1"

/protein_id="BAB18161.1"

/db_xref="GI:11231081"

/translation="MAFHVEGLIAIIVFYLLILLVGIWAAWRTKNSGSAEERSEAIIV
 GGRDIGLLVGGFTMTATWGGGYINGTAEAVYVPGYGLAWAQAPIGYSLSLILGGLFF
 AKPMRSKGYVTMLDPPFQIYGKRMGGLLFIPALMGEMFWAAAIIFSALGATISVIIDVD
 MHISVIISALIATLYTLVGGGLYSVAYTDVVQLFCIFVGLWISVPFALSHPAVADIGFT
 AVHAKYQKPWLGTVDSSSEVYSWLDSFLLMLGGIPWQAYFQVRLSSSSATYAQVLSFL
 AAFGCLVMAIPAILIGAIGASTDWNQTAYGLPDPKTTEEADMILPIVLQYLCPVYISF
 FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVRITVFVFGASA
 TAMALLTKTVYGLWYLSDDLVIYIVIFPQLLCVLFVKGTNTYGAVAGYVSGFLRITGG

EPYLYLQPLIFYPGYYPDDNGIYNQKFPFKTLAMVTSFLTNICISYLAKYLFESGTLF
PKLDVFDVAVARHSEENMDKTIILVKNENIKLDELALVKPRQSMTLSSFTFNKEAFLDV
DSSPEGSGTEDNLQ"

ORIGIN

Query Match 100.0%; Score 1743; DB 9; Length 5158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
          |||
Db      277 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 336

Qy      61 GTTGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
          |||
Db      337 GTTGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 396

Qy     121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
          |||
Db     397 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 456

Qy     181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
          |||
Db     457 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 516

Qy     241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
          |||
Db     517 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 576

Qy     301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
          |||
Db     577 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 636

Qy     361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTTCCTGCACTGATGGGAGAAATGTTT 420
          |||
Db     637 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTTCCTGCACTGATGGGAGAAATGTTT 696

Qy     421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
          |||
Db     697 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 756

Qy     481 ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
          |||
Db     757 ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 816

Qy     541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 600
          |||
Db     817 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 876

Qy     601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
          |||
Db     877 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 936

Qy     661 CATGCCAAATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG 720
          |||
Db     937 CATGCCAAATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG 996
```

Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	780
Db	997	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	1056
Qy	781	GTTCTCTCTTCTTCCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTTCGGG	840
Db	1057	GTTCTCTCTTCTTCCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTTCGGG	1116
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1117	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	1176
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1177	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	1236
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	1237	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1296
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATGTTTGCA	1080
Db	1297	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATGTTTGCA	1356
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1357	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1416
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1417	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1476
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1477	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1536
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1537	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1596
Qy	1321	TCTGGCCTCTTCTGAGAAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1597	TCTGGCCTCTTCTGAGAAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1656
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1657	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1716
Qy	1441	ACACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1717	ACACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1776
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1777	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1836
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620

```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1837 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1896

Qy      1621 GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1897 GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1956

Qy      1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1957 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 2016

Qy      1741 TGA 1743
      |||
Db      2017 TGA 2019

```

RESULT 6

AR268949

LOCUS AR268949 1743 bp DNA linear PAT 10-APR-2003

DEFINITION Sequence 1 from patent US 6500643.

ACCESSION AR268949

VERSION AR268949.1 GI:29699686

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Wu,D.-H., Gu,Y., Millard,W.J. and He,Y.-J.

TITLE Human high affinity choline transporter

JOURNAL Patent: US 6500643-A 1 31-DEC-2002;

FEATURES Location/Qualifiers

source 1..1743

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 99.7%; Score 1738.2; DB 6; Length 1743;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1740; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy      61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180

Qy      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240

Qy      241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300

```

Db	241		GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301		TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301		TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361		ATCTATGGAACCGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC	420
Db	361		ATCTATGGAACCGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC	420
Qy	421		TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421		TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481		ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481		ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541		CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Db	541		CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Qy	601		ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601		ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661		CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661		CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721		CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721		CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781		GTTCTCTCTTCTTCCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781		GTTCTCTCTTCTTCCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841		TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841		TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901		AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901		AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961		CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961		CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021		TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021		TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Qy	1081		CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140

Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAA	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 7

E49870

LOCUS E49870 1743 bp DNA linear PAT 27-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION E49870

VERSION E49870.1 GI:22554901

KEYWORDS JP 2001136976-A/2.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE      1 (bases 1 to 1743)
AUTHORS        Haga,T. and Okuda,T.
TITLE          High-affinity choline transporter
JOURNAL        Patent: JP 2001136976-A 2 22-MAY-2001;
                SCIENCE & TECH AGENCY
COMMENT        OS   Rattus sp. (rat)
                PN   JP 2001136976-A/2
                PD   22-MAY-2001
                PF   27-DEC-1999 JP 1999368991
                PI   TATSUYA HAGA,TAKASHI OKUDA
                PC   C12N15/09,A01K67/027,A61K38/00,C07K14/47,C07K16/18,C07K19/00,
                PC   C12N5/10,
                PC   C12P21/02,C12P21/08,C12Q1/00,C12N15/00,A61K37/02,C12N5/00 CC
                FH   Key                      Location/Qualifiers
                FT   CDS                      (1). .(1743).

FEATURES             Location/Qualifiers
     source           1. .1743
                     /organism="Rattus sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10118"

ORIGIN

Query Match      80.0%; Score 1394.2; DB 6; Length 1743;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 1525; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||||||| ||||||| |||| || ||| ||||||||| || || |||
Db      1 ATGCCTTTCCATGTAGAAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTTATATTTCTG 60

Qy      61 GTTGGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||||||| ||||||| |||| ||||||||| || | ||||||| |||||||||
Db      61 GTTGGGAATATGGGCTGCATGGAAAACCAAAACAGCGGTAATGCAGAAGAACGCAGCGAA 120

Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTTGGTTGGTGGATTTACCATGACAGCT 180
      ||||||||| ||||||| ||||||| ||||||||| ||||||||| |||||||||
Db      121 GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC 180

Qy      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||||| ||||||| || ||||| || ||||||||| ||||||||| ||||||| |
Db      181 ACCTGGGTTGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT 240

Qy      241 GGCCTAGCTTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||
Db      241 GGTCTAGCTTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG 300

Qy      301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||||||| ||||||| |||| ||||||| ||||||||| ||||||||| ||
Db      301 TTTTTTGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG 360

Qy      361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420
      ||||||||| ||||||| || || ||||| || ||||||||| ||||||| |||||||
Db      361 ATCTATGGAAAGCGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTTC 420

Qy      421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480

```

Db	421		TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT	480
Qy	481		ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481		GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCTCGTGGGAGGG	540
Qy	541		CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Db	541		CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG	600
Qy	601		ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601		ATCAGTGTCCCATTTGCCCTGTCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG	660
Qy	661		CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661		CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720
Qy	721		CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	780
Db	721		CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG	780
Qy	781		GTTCTCTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTTCGGG	840
Db	781		GTCCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCCTTCCTGGCAGCTTTTGGG	840
Qy	841		TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841		TGCCTGGTGATGGCTCTACCAGCCATTTGCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901		AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901		AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	960
Qy	961		CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCACTT	1020
Db	961		CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTCTTTGGGCTTGGTGCTGTT	1020
Qy	1021		TCTGCTGCTGTTATGTGCATCAGCAGATTCTTCCATCTTGTCTAGCAAGTTCCATGTTTGCA	1080
Db	1021		TCTGCTGCTGTCATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1080
Qy	1081		CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1081		CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1140
Qy	1141		ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141		ATGAGGATCACTGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1200
Qy	1201		ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201		ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTACATCATCATCTTCCCACAG	1260
Qy	1261		CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320

PC C12N15/12,C07K14/47,C12Q1/68,C07K19/00,C07K16/18,C12N5/10, PC A61K38/17,

CC

FT CDS (1) . . (1743) .

ORIGIN

Query Match 80.0%; Score 1394.2; DB 6; Length 1743;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 1525; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCACCTTCTAATTTTGCTG	60
Db	1	ATGCCTTTCCATGTAGAAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTTATATTTCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCATGGAAAACCAAAACAGCGGTAATGCAGAAGAACGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC	180
Qy	181	ACCTGGGTTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTTCGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG	360
Qy	361	ATCTATGGA AACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Db	361	ATCTATGGA AACGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCTCGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG	600

Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTG	660
Db	601	ATCAGTGTCCCATTTGCCCTGTCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCAGCCATTTCGATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCACTT	1020
Db	961	CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTTCTTTGGGCTTGGTGCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCTAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTCATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTCAGGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTTTTCCCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440

Qy	1441	ACACTTGCCATGGTTACATCATTCTTAAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATATATTTGATGCTGTTGTCTCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTAAAGCCTCGACAGAGCCTAACCCTCAGTTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA	1740
Qy	1741	TGA 1743	
Db	1741	TGA 1743	

RESULT 9

AB030947

LOCUS	AB030947	4904 bp	mRNA	linear	ROD 03-FEB-2000
-------	----------	---------	------	--------	-----------------

AB030947 4904 bp mRNA linear ROD 03-FEB-2000

ACCESSION AB030947

VERSION AB030947.1 GI:6863033

KEYWORDS choline transporter; high-affinity choline transporter CHT1.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM *Rattus norvegicus*

REFERENCE 1 (sites)

AUTHORS Okuda,T., Haga,T., Kanai,Y., Endou,H., Ishihara,T. and Katsura,I.
TITLE Identification and characterization of the high-affinity choline transporter

MEDLINE 20116099

REFERENCE 2 (bases 1 to 4904)

AUTHORS Okuda, T.

COMMENT Sequence updated (11-Jan-2000).

FEATURES	Location/Qualifiers
----------	---------------------

```
source      1.  .4904
```

CDS

```
/strain="Wistar"
/db_xref="taxon:10116"
/clone="CHT1"
/tissue_type="spinal cord"
/clone_lib="rat spinal cord cDNA library"
/dev_stage="adult"
224..1966
/codon_start=1
/product="high-affinity choline transporter CHT1"
/protein_id="BAA90484.1"
/db_xref="GI:6863034"
/translation="MPFHVEGLVAIILFYLLIFLVGIWAAWKTKNSGNAEERSEAIIV
GGRDIGLLVGGFTMTATWVGGGYINGTAEAVYGPCCGLAWAQAPIGYSLSLILGGLFF
AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAI FSALGATISVIIDVD
VNISVIVSALIAILYTLVGGLYSVAYTDVVQLFCIFIGLWISVPFALSHPAVTDIGFT
AVHAKYQSPWLGTIESVEVYTWLDN FLLMLGGIPWQAYFQRLSSSSATYAQVLSFL
AAFGLVLMALPAICIGAIGASTDWNQTAYGFDPKTKEEADMILPIVLQYLCPVYISF
FGLGAVSAAVMSSADSSILSASSMEFARNIYQLSFRONASDKEIVVMRITVFVFGASA
TAMALLTKTVYGLWYLSSDLVYIIIFPQLLCVLFIKGTNTYGAVAGYIFGLFLRITGG
EPYLYLQPLIFYPGYYPDKNGIYNQRFPFKTL SMVTSFFTNICVSYLAKYLFESGTL P
PKLDIFDAVVS RHSEENMDKTILVRNENIKL NELAPVKPRQSLTLSSTFTNKEALLDV
DSSPEGSGTEDNLQ"
```

ORIGIN

Query Match 80.0%; Score 1394.2; DB 10; Length 4904;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 1525; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTGCTG 60
      ||| ||||| ||||| ||||| ||| || ||| ||||| ||||| ||| |||
Db      224 ATGCCTTTCCATGTAGAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTATATTTCTG 283

Qy      61 GTTGGAAATATGGGCTGCCTGGAGAACCAAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      284 GTTGGAAATATGGGCTGCATGGAAAACCAAAAAACAGCGGTAATGCAGAAGAACGCAGCGAA 343

Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      344 GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTGTTGGTGGTGGTTTACCATGACAGCC 403

Qy      181 ACCTGGGTCGGAGGAGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      404 ACCTGGGTTGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT 463

Qy      241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      464 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG 523

Qy      301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      524 TTTTTTGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG 583

Qy      361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      584 ATCTATGGAAAGCGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTTC 643

Qy      421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
```

Db	644	 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT	703
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	704	 GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCCTCGTGGGAGGG	763
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	764	 CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG	823
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	824	 ATCAGTGTCCCATTTGCCCTGTCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG	883
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	884	 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	943
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	944	 CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG	1003
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGTCTCCTTCCTGGCAGCTTTCGGG	840
Db	1004	 GTCCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGTCTCCTTCCTGGCAGCTTTTGGG	1063
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1064	 TGCCTGGTGATGGCTCTACCAGCCATTTGCATTGGGGCCATTGGAGCCTCCACAGACTGG	1123
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1124	 AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	1183
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	1184	 CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTCTTTGGGCTTGGTGCTGTT	1243
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGCAAGTTCCATGTTTGCA	1080
Db	1244	 TCTGCTGCTGTCATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1303
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1304	 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1363
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1364	 ATGAGGATCACTGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1423
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1424	 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTTACATCATCATCTTCCCACAG	1483
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320

Db	1484	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1543
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1544	TTTGGACTTTTCCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1603
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAA	1440
Db	1604	TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAA	1663
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1664	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCCTATCTAGCCAAGTAT	1723
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1724	CTATTTGAAAGTGGAACCTTGCCTCCAAAATTAGATATATTTGATGCTGTTGTCTCAAGG	1783
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1784	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAT	1843
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1844	GAACTTGCACTTGTAAAGCCTCGACAGAGCCTAACCTCAGTTCAACTTTCACCAATAAA	1903
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1904	GAGGCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA	1963
Qy	1741	TGA	1743
Db	1964	TGA	1966

JOURNAL Submitted (09-JUN-2000) Department of Pharmacology and Center for
Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at

Pierce, Nashville, TN 37232-6420, USA
REFERENCE 3 (bases 1 to 1743)
AUTHORS Apparsundaram,S., Ferguson,S.M. and Blakely,R.D.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2001) Department of Pharmacology and Center for
Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at
Pierce, Nashville, TN 37232-6420, USA
REMARK Sequence update by submitter
COMMENT On Feb 28, 2001 this sequence version replaced gi:11527247.
FEATURES Location/Qualifiers

source 1. .1743
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
CDS 1. .1743
/codon_start=1
/product="sodium and chloride-dependent high-affinity
choline transporter"
/protein_id="AAG36945.2"
/db_xref="GI:13162670"
/translation="MPFHVEGLVAIILFYLLIFLVGIWAAWKTKNSGNPEERSEAIIV
GGRDIGLLVGGFTMTATWVGGGYINGTAEAVYGPCCGLAWAHAPIGYSLSLILGGLFF
AKPMRSKGYVTMLDPFKQIYGKRMGGLLFIPALMGEMFWAAAFSALGATISVIIDVD
VNISVIVSALIAILYTLVGGLYSVAYTDVVQLFCIFIGLWISVPFALSHPAVTDIGFT
AVHAKYQSPWLGTIESVEVYTWLDNFLLMLGGIPWQAYFQRVLSSSSATYAQVLSFL
AAFGCLVMALPAICIGAIGASTDWNQTAYGYDPKTKEEADMILPIVLQYLCPVYISF
FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVRITVLVFGASA
TAMALLTKTVYGLWYLSDDLVIYIIIFPQLLCVLFIKGTNTYGAVAGYIFGLFLRITGG
EPYLYLQPLIFYPGYYSKNGIYNQRFPPKTLMTSFFTNICVSYLAKYLFESGTLF
PKLDVFDVAVARHSEENMDKTIILVRNENIKLNELAPVKPRQSLTSSSTFTNKEALLDV
DSSPEGSGTEDNLQ"

ORIGIN

Query Match 78.9%; Score 1375; DB 10; Length 1743;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGCCTTTCCATGTGGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCATGGAAAACCAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGCTTGGTGGATTACCATGACAGCT	180
Db	121	GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTGGTGGTTTACCATGACAGCC	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300

Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCACTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCAATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCTTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGGCCTGTACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCTTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCTATCCTCAGCCACCTATGCTCAGGTACTGTCTTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCACTCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCACTGCTCCTCAGCTGACTCGTCCATCCTGTGCGGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTTCTTTCAGACAAAATGCATCAGACAAAGAAATGTTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200

```

Db      1141  ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG 1200
Qy      1201  ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260
Db      1201  ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG 1260
Qy      1261  CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320
Db      1261  CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1320
Qy      1321  TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380
Db      1321  TTTGGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1380
Qy      1381  TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440
Db      1381  TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1440
Qy      1441  ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500
Db      1441  ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1500
Qy      1501  CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560
Db      1501  CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG 1560
Qy      1561  CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
Db      1561  CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT 1620
Qy      1621  GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTTACCAATAAA 1680
Db      1621  GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTTACCAATAAG 1680
Qy      1681  GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
Db      1681  GAGGCCCTTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAATTTACAA 1740
Qy      1741  TGA 1743
Db      1741  TGA 1743

```

RESULT 11

E49872

LOCUS E49872 1743 bp DNA linear PAT 27-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION E49872

VERSION E49872.1 GI:22554903

KEYWORDS JP 2001136976-A/4.

SOURCE Mus sp.

ORGANISM Mus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Haga, T. and Okuda, T.

TITLE High-affinity choline transporter
JOURNAL Patent: JP 2001136976-A 4 22-MAY-2001;
SCIENCE & TECH AGENCY
COMMENT OS Mus sp. (mouse)
PN JP 2001136976-A/4
PD 22-MAY-2001
PF 27-DEC-1999 JP 1999368991
PI TATSUYA HAGA, TAKASHI OKUDA
PC C12N15/09, A01K67/027, A61K38/00, C07K14/47, C07K16/18, C07K19/00,
PC C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/00, C12N15/00, A61K37/02, C12N5/00 CC
FH Key Location/Qualifiers
FT CDS (1). .(1743).

FEATURES Location/Qualifiers
source 1. .1743
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"

ORIGIN

Query Match 78.8%; Score 1373.4; DB 6; Length 1743;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
||| ||||| || ||||| ||||| ||| | ||||| || || || |||
Db 1 ATGTCTTTCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 60
Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
||||| ||||| ||| ||||| ||||| |||| | ||||| ||| |||
Db 61 GTTGAATATGGGCTGCATGGAAAACCAAAAACAGCGCAACCCAGAAGAGCACAGTGAA 120
Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGGTTGGTGGTGGTTTACCATGACAGCC 180
Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||
Db 181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGATGGGCCAGGTTGT 240
Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 300
Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
|| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCAACAG 360
Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420
||||| ||||| || || || || || || || || || || || || || || || ||
Db 361 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT 420
Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
||||| ||||| || || || || || || || || || || || || || || || ||
Db 421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT 480
Qy 481 ATGCACATTTCTGTCTATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540

Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGGCCTGTGCACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCACTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTGCATCAGCAGATTCTTCCATCTTGTGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTCGGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380

Db 1321 TTTGGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440
 ||||| ||||| ||||| | ||||| ||||| ||||| |||||

Db 1381 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1440

Qy 1441 ACACCTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500
 || || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1441 ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1561 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTACCAATAAA 1680
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTACCAATAAG 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740

Qy 1741 TGA 1743
 |||

Db 1741 TGA 1743

RESULT 12

BD012720

LOCUS BD012720 1743 bp DNA linear PAT 02-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION BD012720

VERSION BD012720.1 GI:22092909

KEYWORDS WO 0116315-A/4.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Haga,T. and Okuda,T.

TITLE High-affinity choline transporter

JOURNAL Patent: WO 0116315-A 4 08-MAR-2001;
 JAPAN SCIENCE AND TECHNOLOGY CORP,TATSUYA HAGA,TAKASHI OKUDA

COMMENT OS Mus musculus (mouse)

PN WO 0116315-A/4

PD 08-MAR-2001

PF 18-AUG-2000 WO 2000JP005545

PR 27-AUG-1999 JP 99P 240642,27-DEC-1999 JP 99P 368991 PI
 TATSUYA HAGA,TAKASHI OKUDA

PC C12N15/12,C07K14/47,C12Q1/68,C07K19/00,C07K16/18,C12N5/10, PC
 A61K38/17,

PC A61K45/00,A61P25/28,G01N33/53,A01K67/027

CC

	FH	Key	Location/Qualifiers
	FT	CDS	(1)..(1743).

FEATURES	Location/Qualifiers
source	1..1743
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"

ORIGIN

Query Match 78.8%; Score 1373.4; DB 6; Length 1743;
 Best Local Similarity 86.7%; Pred. No. 0;
 Matches 1512; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGTCTTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCATGGAAAACCAAAAACAGCGCAACCCAGAAGAGCACAGTGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCACAG	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC	420
Db	361	ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACAGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGGCCCTGTCACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	660

Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661		
		CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721		
	721	CTTGATAATTTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781		
	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG	840
Qy	841	TGCCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841		
	841	TGCCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901		
	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGAGTT	1020
Db	961		
	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGTTCCATGTTTGCA	1080
Db	1021		
	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTCGGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081		
	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141		
	1141	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTACATCGTTATCTTCCCCAG	1260
Db	1201		
	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261		
	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321		
	1321	TTTGGACTATTTCCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381		
	1381	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441		
	1441	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560

```

      |||
Db      1501 CTATTTGAAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG 1560
Qy      1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
      |||
Db      1561 CACAGTGAAGAGAACATGGACAAGACCATTCCTAGTCAGAAATGAAAATATCAAATTAAAT 1620
Qy      1621 GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
      |||
Db      1621 GAACTTGCACCTGTGAAACCTCGGCAGAGCCTAACCTCAGTTCAACTTTCACCAATAAG 1680
Qy      1681 GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
      |||
Db      1681 GAGGCCCTTCCTTGATGTTGATTCCAGTCCGGAGGGGTCTGGGACTGAAGATAACTTACAA 1740
Qy      1741 TGA 1743
      |||
Db      1741 TGA 1743

```

RESULT 13

AX080443

LOCUS AX080443 4938 bp DNA linear PAT 22-FEB-2001

DEFINITION Sequence 1 from Patent WO0078950.

ACCESSION AX080443

VERSION AX080443.1 GI:13159872

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Sierzega,M. and Albrandt,K.

TITLE Differentially expressed genes in the adipocytes of obese mice

JOURNAL Patent: WO 0078950-A 1 28-DEC-2000;

AMYLIN PHARMACEUTICALS, INC. (US)

FEATURES

Location/Qualifiers

source

1. .4938

/organism="Mus musculus"

/mol_type="unassigned DNA"

/db_xref="taxon:10090"

/note="P4P6B1"

ORIGIN

Query Match 78.8%; Score 1373.4; DB 6; Length 4938;

Best Local Similarity 88.7%; Pred. No. 0;

Matches 1512; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      |||
Db      247 ATGTCCTTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATACTTCTG 306
Qy      61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      |||
Db      307 GTTGGAATATGGGCTGCATGGAAAACCAAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 366
Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180

```


Db	1207	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1266
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTCCATGTTTGCA	1080
Db	1267	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTCGGCGAGTTCTATGTTTGCT	1326
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1327	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1386
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1387	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1446
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1447	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTTACATCATCATCTTCCCACAG	1506
Qy	1261	CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1507	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1566
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1567	TTTGGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1626
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1627	TTCTACCCTGGTTATTACTCTGACAAGAAATGGTATATACAATCAGAGGTTCCCATTTAAA	1686
Qy	1441	ACACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1687	ACTCTCTCCATGGTTACCTCATCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1746
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1747	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	1806
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1807	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT	1866
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1867	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAG	1926
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1927	GAGGCCTTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAACTTACAA	1986
Qy	1741	TGA	1743
Db	1987	TGA	1989

LOCUS MMU401467 1743 bp mRNA linear ROD 16-AUG-2000
 DEFINITION Mus musculus mRNA for high affinity choline transporter (CHT1 gene).
 ACCESSION AJ401467
 VERSION AJ401467.1 GI:9843808
 KEYWORDS CHT1 gene; high affinity choline transporter.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Wieland,A., Bonisch,H. and Bruss,M.
 TITLE Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1743)
 AUTHORS Bruess,M.
 TITLE Direct Submission
 JOURNAL Submitted (14-AUG-2000) Bruess M., University of Bonn, Pharmacology and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY
 FEATURES Location/Qualifiers
 source 1. .1743
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/cJ"
 /db_xref="taxon:10090"
 /tissue_type="brainstem"
 gene 1. .1743
 /gene="CHT1"
 CDS 1. .1743
 /gene="CHT1"
 /function="sodium- and chloride-dependent reuptake of choline"
 /codon_start=1
 /evidence=experimental
 /product="high affinity choline transporter"
 /protein_id="CAC03719.1"
 /db_xref="GI:9843809"
 /db_xref="GOA:Q9ESW5"
 /db_xref="SPTREMBL:Q9ESW5"
 /translation="MSFHVEGLVAIIIFYLLIFLVGIWAAWKTKNSGNPEEHSEAIIV
 GGRDIGLLVGGFTMTATWVGGGYINGTAVAVYGPCCGLAWAQAPIGYSLSLILGGLFF
 AKPMRSKGYVTMLDPFQIYGKRMGGLLFIPALMGEMFWAAAFSALGATISVIIDVD
 VNISVIVSALIAILYTLVGGGLYSVAYTDVVQLFCIFIGLWISVPPFALSHPAVTDIGFT
 AVHAKYQSPWLGTIESVEVYTWLDNFLLMLGGIPWQAYFQRVLSSSSATYAQVLSYL
 AAFGCLVMALPAICIGAIGASTDWNQTAYGYDPKTKEEADMILPIVLQYLCPVYISF
 FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVMRITVLVFGASA
 TAMALLTKTVYGLWYLSSDLVYIIIFPQLLCVLFIKGTNTYGAVAGYIFGLFLRITGG
 EPYLYLQPLIFYPGYYSKNGIYNQRFPKTLSMVTSSFTNICVSYLAKYLFESGTLF
 PKLDVFDVAVARHSEENMDKTIILVRNENIKLNELAPVKPRQSLTSLSTFTNKEALLDV
 DSSPEGSGTEDNLQ"

ORIGIN

Query Match 78.4%; Score 1367; DB 10; Length 1743;
 Best Local Similarity 86.5%; Pred. No. 0;
 Matches 1508; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
 ||| ||||| || ||||| ||||| ||| | ||||| || || || |||
 Db 1 ATGTCTTTCCACGTAGAAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 60

Qy 61 GTTGGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 61 GTTGGAATATGGGCTGCATGGAAAACAAAACAGCGGCAACCCAGAAGAGCACAGTGAA 120

Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
 ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
 Db 121 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC 180

Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
 ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
 Db 181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGTAGCAGTGATGGGCCAGGTTGT 240

Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 241 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 300

Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 301 TTTTGTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCACAG 360

Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAAATGTTTC 420
 ||||| ||||| ||||| || || || || || ||||| ||||| ||||| |||||
 Db 361 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC 420

Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
 ||||| ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| |||||
 Db 421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATCGATGTGGAT 480

Qy 481 ATGCACATTTCTGTCTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
 || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 481 GTGAACATATCGGTCAATTGTCTCCGCACTCATTGCCATTCTTTATACCCCTAGTGGGTGGG 540

Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600
 ||||| ||||| || ||||| || ||||| || ||||| || ||||| ||||| ||||| |||||
 Db 541 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 600

Qy 601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 601 ATCAGTGTCCCTTTTGCCCTGTCACATCCTGCAGTCACCGACATCGGATTACAGCTGTG 660

Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG 720
 ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 661 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG 720

Qy 721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 721 CTTGATAATTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 780

Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTTCGGG 840
 || ||||| ||||| ||||| ||||| || ||||| ||||| ||||| ||||| |||||
 Db 781 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTACCTGGCAGCTTTTGGG 840

Qy 841 TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900

Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTATGTCTCTCAGCTGACTCGTCCATCCTGTCGGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Db	1201	ACTGTGTATGGGCTGTGGTACCTGAGCTCTGACCTTGCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTATTTCCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAA	1440
Db	1381	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATTTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATCTATTTGATGCTGTTGTGCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAAATGAAAATATCAAATTAAAT	1620
Qy	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA	1680
Db	1621	GAACCTTGCACTTGTGAAGCCACGAGAGCCTAACCCTCAGTTCAACTTTACCAATAAG	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740

Db 1681 GAGGCCCTCCTTGATGTTGATTCCAGTCCGGAGGGGTCTGGGACTGAAGATAACTTACAA 1740

Qy 1741 TGA 1743

Db 1741 TGA 1743

RESULT 15

TMA420808

LOCUS TMA420808 2528 bp mRNA linear VRT 27-NOV-2001

DEFINITION Torpedo marmorata mRNA for high affinity choline transporter (CHT1 gene).

ACCESSION AJ420808

VERSION AJ420808.1 GI:17148508

KEYWORDS CHT1 gene; high affinity choline transporter.

SOURCE Torpedo marmorata (marbled electric ray)

ORGANISM Torpedo marmorata

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; Elasmobranchii; Squala; Hypnosquala; Pristiorajea; Batoidea; Torpediniformes; Torpedinoidei; Torpedinidae; Torpedo.

REFERENCE

1

AUTHORS Guermonprez, L., O'Regan, S., Meunier, F.M. and Morot-Gaudry-Talarmain, Y.

TITLE Cyclosporin, FK506 and rapamycin inhibit neuronal choline uptake via calcineurin-dependent and independent mechanisms

JOURNAL Unpublished

REFERENCE

2 (bases 1 to 2528)

AUTHORS O'Regan, S.

TITLE Direct Submission

JOURNAL Submitted (21-NOV-2001) O'Regan S., Neurobiologie Cellulaire et Moleculaire, C.N.R.S., 1 av de la Terrasse, F-91198 Gif-sur-Yvette, FRANCE

FEATURES

source

Location/Qualifiers

1..2528

/organism="Torpedo marmorata"

/mol_type="mRNA"

/db_xref="taxon:7788"

/clone="tH312"

/tissue_type="electric lobe"

/tissue_lib="lambda ZAPII ELL"

gene

1..2528

/gene="CHT1"

CDS

49..1803

/gene="CHT1"

/function="neuronal Na-dependent choline transporter"

/codon_start=1

/evidence=experimental

/product="high affinity choline transporter"

/protein_id="CAD12727.1"

/db_xref="GI:17148509"

/db_xref="GOA:Q8UWF0"

/db_xref="SPTREMBL:Q8UWF0"

/translation="MTVHIDGIVAIVLFYLLILFVGLWAAWKSNTSMEGAMDRSEAI
MIGGRDIGLLVGGFTMTATWVGGGYINGTAEAVYVPGYGLAWAQAPFGYALSLVIGGL
FFAKPMRSRGYVTMLDPFQQMYGKRMGGLLFIPALLGEIFWSAAILLSALGATLSVIVD
ININVSVVSAVIAVLYTLVGGGLYSVAYTDVVQLFCIFGLWISIPFALLNPAVTDII
VTANQEVYQEPWVGNIQSKDSLIIWIDNLLLMLGGIPWQVYFQRVLSASSATYAQVLS

FLAAGFCVLMAPSVLIGAIGTSTDWNQTSYGLPGPIGKNETDMILPIVLQHLCPPI
SFFGLGAVSAAMSSADSSILSASSMFARNIYHLAFRQEASDKEIVWVMRITIFLFGG
AATSMALLAQSIYGLWYLSSDLVYVIFPQLISVLFVKGTNTYGSYAGYIIGFLLRIS
GGEPYLMQPFIIYPGCYLDHSFGDDPVYVQRFPFKTMAMLFSLGNTGVSYLVKYL
VSGILPPKLDFLDSVSKHSKEIMDKTFLMNQDNITLSELVHVNPIHSASVSAALTNK
EAFEDIEPNPELSKSGND"

polyA_signal 2487. .2492
/gene="CHT1"

ORIGIN

Query Match 49.7%; Score 867; DB 5; Length 2528;
Best Local Similarity 69.3%; Pred. No. 2.9e-227;
Matches 1217; Conservative 0; Mismatches 520; Indels 18; Gaps 2;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	49	ATGACCGTTCACATCGATGGGATCGTAGCGATCGTCCTGTTTACTTGTTAATCTTATTT	108
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCA-----GCGCAGAAGAGCGC	114
Db	109	GTTGGATTATGGGCTGCTTGGAAGTAAAAACACGTCATGGAAGGAGCAATGGATCGG	168
Qy	115	AGCGAAGCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATG	174
Db	169	AGTGAAGCTATAATGATTGGGGGAAGAGATATCGGGCTGCTGCTTGGTGGCTTCACAATG	228
Qy	175	ACAGCTACCTGGGTCGGAGGAGGGTATATCAATGGCAGCTGAAGCAGTTTATGTACCA	234
Db	229	ACCGCAACTTGGGTCGGTGGCGGTTATATCAATGGGACAGCAGAGGCGGTTTATGTTCT	288
Qy	235	GGTTATGGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGT	294
Db	289	GGGTACGGCTTGGCCTGGGCGCAGGCTCCCTTCGATACGCACTCAGCCTGCTTATTGGC	348
Qy	295	GGCCTGTTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTT	354
Db	349	GGCTTATTTTTCGCTAAACCCATGCGCTCACGGGGTTACGTGACCATGCTGGACCCGTTT	408
Qy	355	CAGCAAATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAA	414
Db	409	CAACAGATGTACGGTAAACGAATGGGAGGATTGCTCTTCATCCCCGCTCTCTGGGGGAA	468
Qy	415	ATGTTCTGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGAT	474
Db	469	ATCTTCTGGTCTGCAGCCATACTGTCCGCGCTAGGTGCAACTTTAAGCGTGATTGTGGAC	528
Qy	475	GTGGATATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTG	534
Db	529	ATCAATATAAACGTATCAGTGGTAGTTCCGCTGTGATCGCTGTATTATACACTCTGGTC	588
Qy	535	GGAGGGCTCTATTCTGTGGCCTACACTGATGTCGTTTCTTCTGATTTTGTAGGG	594
Db	589	GGCGGGTTATACTCGGTCGCTACACAGATGTCGTCAGTTGTTTTGCATCTTCTTAGGT	648
Qy	595	CTGTGGATCAGCGTCCCCCTTTGCATTGTGCACATCCTGCAGTCGACAGATCGGGTTCACT	654
Db	649	TTGTGGATCAGCATCCCTTTCGCCCTCTTAAATCCCGCTGTTACGGATATTATCGTGACC	708

Qy	655	GCTGTGCATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTTGACTCATCTGAAGTCTAC	714
Db	709	GCAAATCAAGAAGTTTATCAGGAGCCTTGGGTGGGAAATATAACAATCAAAGGACAGTTTA	768
Qy	715	TCTTGGCTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTT	774
Db	769	ATCTGGATTGACAACCTTTCTATTACTGATGCTGGGTGGAATCCCGTGGCAAGTATATTTT	828
Qy	775	CAGAGGGTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCTCCTGGCAGCT	834
Db	829	CAGAGAGTCCTTTCTGCTTCTTCTGCTACCTATGCGCAAGTCCTGTCTTTCTGGCTGCC	888
Qy	835	TTCGGGTGCCTGGTGATGGCCATCCCAGCCATACTCATTTGGGGCCATTGGAGCATCAACA	894
Db	889	TTCGGTTGCGTTCTCATGGCCATCCCGTCTGTTCATCGGTGCAATAGGAACATCCACT	948
Qy	895	GACTGGAACCAGACTGCATATGGGCTTCCAGATCCAAGACTACAGAAGAGGCAGACATG	954
Db	949	GACTGGAATCAGACTTCCTATGGCTTGCCAGGCCCTATAGGCCAAAATGAGACTGATATG	1008
Qy	955	ATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGT	1014
Db	1009	ATTTTGCCGATCGTGCTGCAGCATCTGTGTCCACCCTACATTTCTTTTGGTCTTGGC	1068
Qy	1015	GCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATG	1074
Db	1069	GCTGTCTCTGCTGCTGTGATGTCATCGGCTGATTCTTCTATCTTATCAGCAAGTTCTATG	1128
Qy	1075	TTTGCACGGAACATCTACCAGCTTTCTTCAGACAAAATGCTTCGGACAAAGAAATCGTT	1134
Db	1129	TTTGCTCGGAATATTTACCATCTTGCTTTCAGACAAGAGGCTTCAGACAAAGAAATAGTG	1188
Qy	1135	TGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTG	1194
Db	1189	TGGGTAATGCGAATCACCATATTTCTATTTGGAGGAGCTGCAACATCTATGGCATTGCTT	1248
Qy	1195	ACGAAAACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTC	1254
Db	1249	GCTCAATCAATCTATGGCCTCTGGTATCTGAGCTCAGATCTTGTCTACGTCATTATCTTT	1308
Qy	1255	CCCCAGCTGCTTTGTGTA CTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGT	1314
Db	1309	CCTCAATTAATATCAGTGCTCTTCGTCAAGGGAACAAACACATATGGGTCTATTGCTGGA	1368
Qy	1315	TATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCC	1374
Db	1369	TATATCATTTGGCTTTTTGCTTCGGATTAGTGGTGGTGAACCATATTTACATATGCAGCCA	1428
Qy	1375	TTGATCTTCTACCCTGGCT-----ATTACCCTGATGATAATGGTATATATAAT	1422
Db	1429	TTTATTTATTACCCTGGATGCTATTTAGATCATTCCTTTGGAGATGATCCTGTTTATGTT	1488
Qy	1423	CAGAAATTTCCATTTAAAAACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATC	1482
Db	1489	CAGAGATTTCCCTTTAAAAACCATGGCAATGTTATTCTCCTTCTTGGGCAACACTGGTGTA	1548

Qy	1483	TCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTT	1542
Db	1549	TCATATCTTGTCAAGTACCTGTTTCGTAAGTGGAATATTGCCACCAAAATTAGACTTCCTT	1608
Qy	1543	GATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAAT	1602
Db	1609	GACAGCGTTGTATCAAAACACAGTAAGGAAATCATGGACAAAACATTCTTGATGAATCAG	1668
Qy	1603	GAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGC	1662
Db	1669	GACAATATTACTTTGTCAGAGCTGGTGCATGTTAATCCAATACACAGTGCTTCAGTTAGT	1728
Qy	1663	TCAACTTTCACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGG	1722
Db	1729	GCTGCTTTGACCAATAAGGAAGCATTGGAAGACATTGAGCCAAATCCTGAACTTTCTAAG	1788
Qy	1723	ACTGAAGATAATTTA	1737
Db	1789	TCAGGCAATGATTGA	1803

Search completed: September 29, 2004, 00:10:03
Job time : 10880 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2004, 19:36:56 ; Search time 1009 Seconds
(without alignments)
7338.562 Million cell updates/sec

Title: US-10-069-541-5
Perfect score: 1743
Sequence: 1 atggctttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1743	100.0	1743	4	AAF81712	Aaf81712 Human hig
2	1743	100.0	1743	5	AAH49207	Aah49207 Human CHO
3	1743	100.0	1743	9	ADD50638	Add50638 cDNA enco
4	1743	100.0	1813	9	ADD50646	Add50646 High-affi
5	1738.2	99.7	1743	8	ABX94338	Abx94338 Human cDN
6	1394.2	80.0	1743	4	AAF81711	Aaf81711 Rat high
7	1394.2	80.0	4904	9	ADD50642	Add50642 cDNA enco

8	1375	78.9	1743	9	ADD50640	Add50640 cDNA enco
9	1375	78.9	1743	9	ADD50660	Add50660 cDNA enco
10	1373.4	78.8	1743	4	AAF81713	Aaf81713 Mouse hig
11	1373.4	78.8	4938	5	AAD02457	Aad02457 Mouse P4P
12	630.8	36.2	8760	5	AAH49206	Aah49206 Human CHO
13	630.8	36.2	119040	9	ADD50656	Add50656 BAC seque
14	630.8	36.2	142299	9	ADD50651	Add50651 BAC seque
15	363.8	20.9	1731	4	AAF81710	Aaf81710 C. elegan
16	363.8	20.9	1985	9	ADD50644	Add50644 High-affi
17	279.6	16.0	386	5	AAD02461	Aad02461 Mouse P4P
18	266.6	15.3	1729	4	ABL29569	Abl29569 Drosophil
19	242.6	13.9	1461	6	ABK73210	Abk73210 Bacillus
20	180.8	10.4	10140	5	AAH49201	Aah49201 Human CHO
c 21	180.8	10.4	119040	9	ADD50656	Add50656 BAC seque
c 22	180.8	10.4	142299	9	ADD50651	Add50651 BAC seque
c 23	179.6	10.3	4223	4	ABL29568	Abl29568 Drosophil
24	167.6	9.6	1094	5	AAS82193	Aas82193 DNA encod
25	163	9.4	240	5	AAH49203	Aah49203 Human CHO
26	155	8.9	455	4	AAI11913	Aai11913 Probe #18
27	155	8.9	455	4	ABA53620	Aba53620 Human foe
28	155	8.9	455	4	AAI33245	Aai33245 Probe #19
29	155	8.9	455	4	ABA23372	Aba23372 Probe #18
30	155	8.9	455	4	AAK27341	Aak27341 Human bon
31	155	8.9	455	4	AAK01883	Aak01883 Human bra
32	155	8.9	455	4	ABS26913	Abs26913 Human liv
33	155	8.9	455	5	AAI01851	Aai01851 Probe #18
34	154.6	8.9	300	5	AAH49204	Aah49204 Human CHO
35	150.8	8.7	240	5	AAH49205	Aah49205 Human CHO
36	108	6.2	180	5	AAH49202	Aah49202 Human CHO
37	72	4.1	96	4	AAI21121	Aai21121 Probe #11
38	72	4.1	96	4	ABA66205	Aba66205 Human foe
39	72	4.1	96	4	AAI46388	Aai46388 Probe #15
40	72	4.1	96	4	ABA33269	Aba33269 Probe #11
41	72	4.1	96	4	AAK40366	Aak40366 Human bon
42	72	4.1	96	4	AAK14621	Aak14621 Human bra
43	72	4.1	96	4	ABS39934	Abs39934 Human liv
44	72	4.1	96	5	AAI06841	Aai06841 Probe #68
45	60	3.4	60	6	ABN37501	Abn37501 Human spl

ALIGNMENTS

RESULT 1

AAF81712

ID AAF81712 standard; cDNA; 1743 BP.

XX

AC AAF81712;

XX

DT 01-JUN-2001 (first entry)

XX

DE Human high affinity choline transporter protein encoding cDNA.

XX

KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
KW ss.

XX

OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT CDS 1. .1743
 FT /*tag= a
 FT /product= "high affinity choline transporter"
 XX
 PN WO200116315-A1.
 XX
 PD 08-MAR-2001.
 XX
 PF 18-AUG-2000; 2000WO-JP005545.
 XX
 PR 27-AUG-1999; 99JP-00240642.
 PR 27-DEC-1999; 99JP-00368991.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Haga T, Okuda T;
 XX
 DR WPI; 2001-226688/23.
 DR P-PSDB; AAB74665.
 XX
 PT New rat and human spinal cord high affinity choline transporters, useful
 PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
 PT treating Alzheimer's disease.
 XX
 PS Claim 9; Page 71-75; 90pp; Japanese.
 XX
 CC The present sequence encodes a human (Homo sapiens) high affinity choline
 CC transporter protein designated cho-1. The cho-1 protein has nootropic and
 CC neuroprotective activities. The cho-1 polynucleotide and protein can be
 CC used for the diagnosis of diseases related to the expression of cho-1 by
 CC comparing the cho-1 polynucleotide sequence in a sample to that of a
 CC control. Drug compositions containing the cho-1 protein or expression
 CC promoters or inhibitors of cho-1 are useful for treating disorders
 CC characterised by abnormal levels of cho-1, such as Alzheimer's disease
 XX
 SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 4; Length 1743;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240

Db	181	ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Db	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCAATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCAATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080

Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 2

AAH49207

ID AAH49207 standard; cDNA; 1743 BP.

XX

AC AAH49207;

XX

DT 26-NOV-2001 (first entry)

XX
 DE Human CHOT encoding cDNA.
 XX
 KW CHOT; human; choline transporter; chromosome 2q11-13; nootropic;
 KW neuroprotective; gene therapy; antisense therapy; degenerative disease;
 KW cognitive disorder; Alzheimer's disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN DE10009055-A1.
 XX
 PD 30-AUG-2001.
 XX
 PF 28-FEB-2000; 2000DE-01009055.
 XX
 PR 28-FEB-2000; 2000DE-01009055.
 XX
 PA (BRUE/) BRUESS M.
 PA (BOEN/) BOENISCH H.
 XX
 PI Bruess M, Boenisch H;
 XX
 DR WPI; 2001-590709/67.
 DR P-PSDB; AAB86837.
 XX
 PT A new gene encoding human choline transporter, designated hCHOT is
 PT located on chromosome 2q11-13 and is useful to treat degenerative
 PT disorders such as Alzheimer's disease.
 XX
 PS Disclosure; Page 11; 12pp; German.
 XX
 CC This invention describes a novel gene encoding human choline transporter,
 CC designated hCHOT which is located on chromosome 2q11-13. The products of
 CC the invention have nootropic and neuroprotective activity and can be used
 CC for gene or antisense therapy. (I) is used to treat degenerative disease,
 CC particularly cognitive disorders such as Alzheimer's disease. Sense and
 CC antisense oligonucleotides derived from the gene may be used in
 CC diagnostics and other techniques. This sequence encodes the human CHOT
 CC protein described in the invention
 XX
 SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 5; Length 1743;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCACAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCCTGGAGAACCACAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180

Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAACCGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAACCGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020

Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAAATATTAAATTAGAT	1620
Qy	1621	GAACCTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAAA	1680
Db	1621	GAACCTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAAA	1680
Qy	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 3

ADD50638

ID ADD50638 standard; cDNA; 1743 BP.

XX

AC ADD50638;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE cDNA encoding human high-affinity choline transporter (hCHT).
 XX
 KW Human; high-affinity choline transporter; hCHT; chromosome 2q12;
 KW cholinergic function; Parkinson's disease; Huntington's disease;
 KW Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;
 KW brain; cholinergic signalling; antiparkinsonian; anticonvulsant;
 KW nootropic; neuroprotective; neuroleptic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1. .1743
 FT /*tag= a
 FT /product= "hCHT"
 XX
 PN US2003114399-A1.
 XX
 PD 19-JUN-2003.
 XX
 PF 23-JUL-2001; 2001US-00911077.
 XX
 PR 23-JUL-2001; 2001US-00911077.
 XX
 PA (BLAK/) BLAKELY R D.
 PA (APPA/) APPARSUNDARAM S.
 PA (FERG/) FERGUSON S.
 XX
 PI Blakely RD, Apparsundaram S, Ferguson S;
 XX
 DR WPI; 2003-810914/76.
 DR P-PSDB; ADD50639.
 XX
 PT Novel isolated polynucleotide encoding human or mouse high affinity
 PT choline transporter polypeptide, useful in gene therapy to increase
 PT cholinergic function in a cell of a patient suffering from Alzheimer's
 PT disease.
 XX
 PS Claim 2; SEQ ID NO 1; 74pp; English.
 XX
 CC The present invention relates to the isolation of polynucleotide
 CC sequences encoding human and mouse high-affinity choline transporter
 CC (hCHT and mCHT respectively), and the proteins they encode. The gene
 CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
 CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT
 CC polynucleotide sequence when delivered to a cell, increases cholinergic
 CC function in the cell that is in a patient having Parkinson's disease,
 CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
 CC myasthenia gravis. The hCHT antibody is useful for controlling
 CC transporter CHT proteins to the brain, and for treating the above
 CC mentioned diseases. The antibody is also useful for diagnosing the above
 CC mentioned disorders and to detect the influence of cholinergic
 CC signalling. The present sequence encodes hCHT. Note: The sequence data
 CC for this patent was obtained in electronic format directly from the USPTO

CC web site at seqdata.uspto.gov.

XX

SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCACAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCCTGGAGAACCACAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780

Db	721	 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCATGTTTGCA	1080
Db	1021	 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAAATGAAAAATATTAAATTAGAT	1620

Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
 QY 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCCACCAATAAA 1680
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCCACCAATAAA 1680
 QY 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 QY 1741 TGA 1743
 |||
 Db 1741 TGA 1743

RESULT 4

ADD50646

ID ADD50646 standard; DNA; 1813 BP.

XX

AC ADD50646;

XX

DT 15-JAN-2004 (first entry)

XX

DE High-affinity choline transporter (CHT) associated DNA sequence #2.

XX

KW High-affinity choline transporter; CHT; cholinergic function;

KW Parkinson's disease; Huntington's disease; Alzheimer's disease;

KW schizophrenia; dysautonomia; myasthenia gravis; brain;

KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;

KW neuroprotective; neuroleptic; ds.

XX

OS Unidentified.

XX

PN US2003114399-A1.

XX

PD 19-JUN-2003.

XX

PF 23-JUL-2001; 2001US-00911077.

XX

PR 23-JUL-2001; 2001US-00911077.

XX

PA (BLAK/) BLAKELY R D.

PA (APPA/) APPARSUNDARAM S.

PA (FERG/) FERGUSON S.

XX

PI Blakely RD, Apparsundaram S, Ferguson S;

XX

DR WPI; 2003-810914/76.

XX

PT Novel isolated polynucleotide encoding human or mouse high affinity

PT choline transporter polypeptide, useful in gene therapy to increase

PT cholinergic function in a cell of a patient suffering from Alzheimer's

PT disease.

XX

PS Disclosure; SEQ ID NO 9; 74pp; English.

XX

CC The present invention relates to the isolation of polynucleotide

CC sequences encoding human and mouse high-affinity choline transporter
 CC (hCHT and mCHT respectively), and the proteins they encode. The gene
 CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
 CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT
 CC polynucleotide sequence when delivered to a cell, increases cholinergic
 CC function in the cell that is in a patient having Parkinson's disease,
 CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
 CC myasthenia gravis. The hCHT antibody is useful for controlling
 CC transporter CHT proteins to the brain, and for treating the above
 CC mentioned diseases. The antibody is also useful for diagnosing the above
 CC mentioned disorders and to detect the influence of cholinergic
 CC signalling. The present DNA sequence of unknown function is provided in
 CC the electronic sequence data but is not mentioned in the printed
 CC specification. Note: The sequence data for this patent was obtained in
 CC electronic format directly from the USPTO web site at seqdata.uspto.gov.
 XX
 SQ Sequence 1813 BP; 440 A; 406 C; 417 G; 550 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 9; Length 1813;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	19	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	78
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACC AAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	79	GTTGGAATATGGGCTGCCTGGAGAACC AAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	138
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Db	139	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	198
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	199	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	258
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	259	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	318
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	319	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	378
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Db	379	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	438
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	439	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	498
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	499	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	558

Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG	600
Db	559	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG	618
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	619	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	678
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	679	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	738
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	739	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	798
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGTCTCCTTCCTGGCAGCTTTCGGG	840
Db	799	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGTCTCCTTCCTGGCAGCTTTCGGG	858
Qy	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	859	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	918
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	919	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	978
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	979	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1038
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1039	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1098
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1099	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1158
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1159	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1218
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1219	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1278
Qy	1261	CTGCTTTGTGTACTCTTTGTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1279	CTGCTTTGTGTACTCTTTGTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1338
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1339	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1398

Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1399	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1458
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1459	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1518
Qy	1501	CTATTTGAAAGTGGAAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1519	CTATTTGAAAGTGGAAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1578
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1579	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1638
Qy	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1639	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1698
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1699	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1758
Qy	1741	TGA	1743
Db	1759	TGA	1761

RESULT 5

ABX94338

ID ABX94338 standard; cDNA; 1743 BP.

XX

AC ABX94338;

XX

DT 13-JUN-2003 (first entry)

XX

DE Human cDNA encoding high affinity choline transporter, HACT.

XX

KW Human; ss; gene; HACT; high affinity choline transporter; pain;
KW neurotransmitter biosynthesis; learning and memory; aging; epilepsy;
KW neurological disorder; spasticity; myoclonus; muscle spasm;
KW muscle hyperactivity; stroke; head trauma; neuronal cell death;
KW multiple sclerosis; spinal chord injury; dystonia; Alzheimer's disease;
KW Myasthenia Gravis; multi-infarct dementia; AIDS dementia;
KW Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW ALS; attention deficit disorder; organic brain syndrome; schizophrenia;
KW nicotine addiction; memory disorder; cognitive disorder.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1. .1743

FT /*tag= a

FT /product= "HACT"

XX

PN US6500643-B1.

XX
 PD 31-DEC-2002.
 XX
 PF 07-SEP-2000; 2000US-00657252.
 XX
 PR 07-SEP-2000; 2000US-00657252.
 XX
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Wu D, Gu Y, Millard WJ, He Y;
 XX
 DR WPI; 2003-361535/34.
 DR P-PSDB; ABU08979.
 XX
 PT Novel isolated polynucleotide (I) that encodes high affinity choline
 PT transporter protein, useful for preventing, treating or ameliorating
 PT neurological and cognitive disorders such as Alzheimer's or Parkinson's
 PT disease.
 XX
 PS Claim 2; Col 17-21; 20pp; English.
 XX
 CC The invention relates to an isolated polynucleotide which encodes a high
 CC affinity choline transporter (HACT) protein appearing as ABU08979. Also
 CC included are a polynucleotide encoding a fragment consisting of at least
 CC about 50 amino acids of the HACT protein, a vector comprising the
 CC polynucleotide, a composition comprising a vector comprising a
 CC polynucleotide which comprises at least about 12 contiguous nucleic acids
 CC of a polynucleotide appearing as ABX94339 (encoding choline
 CC acetyltransferase), a recombinant host cell which comprises the vector
 CC (used to express the HACT protein or fragment). The polynucleotide is
 CC useful as a probe or primer to detect the presence of HACT polynucleotide
 CC in a sample, such as a biological sample, or for screening for test
 CC agents which bind to the polynucleotide. A pharmaceutical composition
 CC comprising the polynucleotide is useful for preventing, treating or
 CC ameliorating neurological and cognitive disorders e.g. pain, spasticity,
 CC myoclonus, muscle spasm, muscle hyperactivity, epilepsy, stroke, head
 CC trauma, neuronal cell death, multiple sclerosis, spinal chord injury,
 CC dystonia, Alzheimer's disease, myasthenia gravis, multi-infarct
 CC dementia, AIDS dementia, Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis (ALS), attention deficit disorder, nicotine
 CC addiction, organic brain syndromes, schizophrenia or memory and cognitive
 CC disorders. HACT is thought to be the rate limiting step in cholinergic
 CC neurotransmitter biosynthesis and regeneration (cholinergic transmissions
 CC are crucial to brain functions such as learning and memory). The present
 CC sequence encodes human HACT
 XX
 SQ Sequence 1743 BP; 411 A; 395 C; 405 G; 532 T; 0 U; 0 Other;

 Query Match 99.7%; Score 1738.2; DB 8; Length 1743;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1740; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

 Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

 Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Db	61		61 GTTGGAAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121		121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121		121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Qy	181		181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181		181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241		241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241		241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301		301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301		301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361		361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Db	361		361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Qy	421		421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421		421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481		481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481		481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541		541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Db	541		541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Qy	601		601 ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601		601 ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661		661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661		661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721		721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	780
Db	721		721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	780
Qy	781		781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781		781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841		841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841		841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901		901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960

Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAA	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 6

AAF81711

ID AAF81711 standard; cDNA; 1743 BP.

XX

AC AAF81711;

XX

DT 01-JUN-2001 (first entry)

XX

DE Rat high affinity choline transporter protein encoding cDNA.

XX

KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
KW ss.

XX

OS Rattus norvegicus.

XX

FH Key Location/Qualifiers

FT CDS 1. .1743

FT /*tag= a

FT /product= "high affinity choline transporter"

XX

PN WO200116315-A1.

XX

PD 08-MAR-2001.

XX

PF 18-AUG-2000; 2000WO-JP005545.

XX

PR 27-AUG-1999; 99JP-00240642.

PR 27-DEC-1999; 99JP-00368991.

XX

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX

PI Haga T, Okuda T;

XX

DR WPI; 2001-226688/23.

DR P-PSDB; AAB74664.

XX

PT New rat and human spinal cord high affinity choline transporters, useful
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT treating Alzheimer's disease.

XX

PS Claim 6; Page 64-68; 90pp; Japanese.

XX

CC The present sequence encodes a rat (Rattus norvegicus) high affinity
CC choline transporter protein designated cho-1. The cho-1 protein has
CC nootropic and neuroprotective activities. The cho-1 polynucleotide and
CC protein can be used for the diagnosis of diseases related to the
CC expression of cho-1 by comparing the cho-1 polynucleotide sequence in a
CC sample to that of a control. Drug compositions containing the cho-1
CC protein or expression promoters or inhibitors of cho-1 are useful for
CC treating disorders characterised by abnormal levels of cho-1, such as
CC Alzheimer's disease

XX

SQ Sequence 1743 BP; 414 A; 402 C; 404 G; 523 T; 0 U; 0 Other;

Query Match

80.0%; Score 1394.2; DB 4; Length 1743;

Db	781	GTCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCTTTCCTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCAGCCATTTGCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAAACCTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTTCTTTGGGCTTGGTGTCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTCTATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTCTACATCATCATCTTCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTTTTTCCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGACGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATTTCTTACCAACATTTGTGTTTCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATATATTTGATGCTGTTGTCTCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAAT	1620
Qy	1621	GAACCTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACCTGCACCTGTAAAGCCTCGACAGAGCCTAACCCCTCAGTTCAACTTTCACCAATAAA	1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 ||||| |||||
 Db 1681 GAGGCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA 1740
 |||||
 Qy 1741 TGA 1743
 |||||
 Db 1741 TGA 1743

RESULT 7

ADD50642

ID ADD50642 standard; cDNA; 4904 BP.

XX

AC ADD50642;

XX

DT 15-JAN-2004 (first entry)

XX

DE cDNA encoding rat high-affinity choline transporter (rCHT).

XX

KW Rat; high-affinity choline transporter; rCHT; cholinergic function;

KW Parkinson's disease; Huntington's disease; Alzheimer's disease;

KW schizophrenia; dysautonomia; myasthenia gravis; brain;

KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;

KW neuroprotective; neuroleptic; gene; ss.

XX

OS Rattus sp.

XX

FH Key Location/Qualifiers

FT CDS 224..1966

FT /*tag= a

FT /product= "rCHT"

XX

PN US2003114399-A1.

XX

PD 19-JUN-2003.

XX

PF 23-JUL-2001; 2001US-00911077.

XX

PR 23-JUL-2001; 2001US-00911077.

XX

PA (BLAK/) BLAKELY R D.

PA (APPA/) APPARSUNDARAM S.

PA (FERG/) FERGUSON S.

XX

PI Blakely RD, Apparsundaram S, Ferguson S;

XX

DR WPI; 2003-810914/76.

DR P-PSDB; ADD50643.

XX

PT Novel isolated polynucleotide encoding human or mouse high affinity

PT choline transporter polypeptide, useful in gene therapy to increase

PT cholinergic function in a cell of a patient suffering from Alzheimer's

PT disease.

XX

PS Example 1; SEQ ID NO 5; 74pp; English.

XX

Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG 600
 ||||| ||||| || ||||| || ||||| || ||||| ||||| |||||
 Db 764 CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTTATAGGATTGTGG 823

Qy 601 ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTGTG 660
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 824 ATCAGTGTCCCATTTGCCCTGTACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG 883

Qy 661 CATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720
 ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 884 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG 943

Qy 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 944 CTTGATAATTTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG 1003

Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840
 || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1004 GTCCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCCTTCCTGGCAGCTTTCGGG 1063

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900
 ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 1064 TGCCTGGTGATGGCTCTACCAGCCATTTGCATTGGGGCCATTGGAGCCTCCACAGACTGG 1123

Qy 901 AACCAGACTGCATATGGGCTTCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960
 ||||| ||||| ||||| ||||| ||||| ||||| || ||||| ||||| |||||
 Db 1124 AACCAAACGCATATGGGTTTCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC 1183

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCA 1020
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1184 CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTCTTTGGGCTTGGTGCTGTT 1243

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGCAAGTTCCATGTTTGCA 1080
 ||||| ||||| ||||| || || || || || || ||||| ||||| |||||
 Db 1244 TCTGCTGCTGTCATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT 1303

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140
 ||||| ||||| ||||| ||||| ||||| ||||| || ||||| ||||| |||||
 Db 1304 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC 1363

Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200
 ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1364 ATGAGGATCACTGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG 1423

Qy 1201 ACTGTGTATGCGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260
 ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 1424 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTACATCATCTTCCACAG 1483

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320
 ||||| ||||| ||||| || ||||| ||||| ||||| || ||||| |||||
 Db 1484 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1543

Qy 1321 TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380
 | ||| || ||||| || ||||| ||||| || || ||||| ||||| |||||
 Db 1544 TTTGGACTTTTCCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1603

XX 23-JUL-2001; 2001US-00911077.
XX
PA (BLAK/) BLAKELY R D.
PA (APPA/) APPARSUNDARAM S.
PA (FERG/) FERGUSON S.
XX
PI Blakely RD, Apparsundaram S, Ferguson S;
XX
DR WPI; 2003-810914/76.
DR P-PSDB; ADD50641.
XX
PT Novel isolated polynucleotide encoding human or mouse high affinity
PT choline transporter polypeptide, useful in gene therapy to increase
PT cholinergic function in a cell of a patient suffering from Alzheimer's
PT disease.
XX
PS Claim 30; SEQ ID NO 3; 74pp; English.
XX
CC The present invention relates to the isolation of polynucleotide
CC sequences encoding human and mouse high-affinity choline transporter
CC (hCHT and mCHT respectively), and the proteins they encode. The gene
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT
CC polynucleotide sequence when delivered to a cell, increases cholinergic
CC function in the cell that is in a patient having Parkinson's disease,
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
CC myasthenia gravis. The hCHT antibody is useful for controlling
CC transporter CHT proteins to the brain, and for treating the above
CC mentioned diseases. The antibody is also useful for diagnosing the above
CC mentioned disorders and to detect the influence of cholinergic
CC signalling. The present sequence encodes mCHT. Note: The sequence data
CC for this patent was obtained in electronic format directly from the USPTO
CC web site at seqdata.uspto.gov.
XX
SQ Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

Query Match 78.9%; Score 1375; DB 9; Length 1743;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGCCTTTCCATGTGGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCATGGAAAACCAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTTGGTTGGTGGTTTTACCATGACAGCC	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	240

Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Db	361	ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCACTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCAATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTGCCCTGTACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTAAGTGTCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCACTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCACTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCTCAGCTGACTCGTCCATCCTGTGCGGAGATTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTTCTTCCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140

XX
 KW Mouse; high-affinity choline transporter; mCHT; cholinergic function;
 KW Parkinson's disease; Huntington's disease; Alzheimer's disease;
 KW schizophrenia; dysautonomia; myasthenia gravis; brain;
 KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;
 KW neuroprotective; neuroleptic; gene; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1. .1743
 FT /*tag= a
 FT /product= "mCHT #2"
 XX
 PN US2003114399-A1.
 XX
 PD 19-JUN-2003.
 XX
 PF 23-JUL-2001; 2001US-00911077.
 XX
 PR 23-JUL-2001; 2001US-00911077.
 XX
 PA (BLAK/) BLAKELY R D.
 PA (APPA/) APPARSUNDARAM S.
 PA (FERG/) FERGUSON S.
 XX
 PI Blakely RD, Apparsundaram S, Ferguson S;
 XX
 DR WPI; 2003-810914/76.
 DR P-PSDB; ADD50661.
 XX
 PT Novel isolated polynucleotide encoding human or mouse high affinity
 PT choline transporter polypeptide, useful in gene therapy to increase
 PT cholinergic function in a cell of a patient suffering from Alzheimer's
 PT disease.
 XX
 PS Example 4; SEQ ID NO 23; 74pp; English.
 XX
 CC The present invention relates to the isolation of polynucleotide
 CC sequences encoding human and mouse high-affinity choline transporter
 CC (hCHT and mCHT respectively), and the proteins they encode. The gene
 CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
 CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT
 CC polynucleotide sequence when delivered to a cell, increases cholinergic
 CC function in the cell that is in a patient having Parkinson's disease,
 CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
 CC myasthenia gravis. The hCHT antibody is useful for controlling
 CC transporter CHT proteins to the brain, and for treating the above
 CC mentioned diseases. The antibody is also useful for diagnosing the above
 CC mentioned disorders and to detect the influence of cholinergic
 CC signalling. The present sequence encodes mCHT. Note: The sequence data
 CC for this patent was obtained in electronic format directly from the USPTO
 CC web site at seqdata.uspto.gov.
 XX
 SQ Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

Query Match

78.9%; Score 1375; DB 9; Length 1743;

Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1 ATGCCTTTCCATGTGGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTCTG 60

Qy     61 GTTGGAAATATGGGCTGCCTGGAGAACCACAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     61 GTTGGAAATATGGGCTGCATGGAAAACCACAAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    121 GCCATCATAGTCGGGGGCGCTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC 180

Qy    181 ACCTGGGTTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    181 ACCTGGGTTCGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    241 GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 300

Qy    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    301 TTTTTCGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG 360

Qy    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    361 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT 480

Qy    481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    481 GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG 540

Qy    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTAGCTCTTTTGCATTTTGTAGGGCTGTGG 600
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    541 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 600

Qy    601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    601 ATCAGTGTCCCCTTTTGCCCTGTCACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG 660

Qy    661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    661 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG 720

Qy    721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    721 CTTGATAATTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 780

Qy    781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTCGGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTGTGTCAGGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGTCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTATTACTCTGACAAGAAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCAFTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCAFTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCTTCAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAAATATCAAATTAAAT	1620
Qy	1621	GAACCTGCACTTGTTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACCTGCACCTGTGAAACCTCGGCAGAGCCTAACCCCTCAGTTCAACTTTCACCAATAAG	1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 ||||| |||||
 Db 1681 GAGGCCCTCCTTGATGTTGATTCCAGTCCGGAGGGGTCTGGGACTGAAGATAATTTACAA 1740
 Qy 1741 TGA 1743
 |||
 Db 1741 TGA 1743

RESULT 10

AAF81713

ID AAF81713 standard; cDNA; 1743 BP.

XX

AC AAF81713;

XX

DT 01-JUN-2001 (first entry)

XX

DE Mouse high affinity choline transporter protein encoding cDNA.

XX

KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;

KW ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS

1. .1743

FT

/*tag= a

FT

/product= "high affinity choline transporter"

XX

PN WO200116315-A1.

XX

PD 08-MAR-2001.

XX

PF 18-AUG-2000; 2000WO-JP005545.

XX

PR 27-AUG-1999; 99JP-00240642.

PR

27-DEC-1999; 99JP-00368991.

XX

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX

PI Haga T, Okuda T;

XX

DR WPI; 2001-226688/23.

DR

P-PSDB; AAB74666.

XX

PT New rat and human spinal cord high affinity choline transporters, useful
 PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
 PT treating Alzheimer's disease.

XX

PS Claim 12; Page 78-82; 90pp; Japanese.

XX

CC The present sequence encodes a mouse (Mus musculus) high affinity choline
 CC transporter protein designated cho-1. The cho-1 protein has nootropic and
 CC neuroprotective activities. The cho-1 polynucleotide and protein can be
 CC used for the diagnosis of diseases related to the expression of cho-1 by
 CC comparing the cho-1 polynucleotide sequence in a sample to that of a

CC control. Drug compositions containing the cho-1 protein or expression
CC promoters or inhibitors of cho-1 are useful for treating disorders
CC characterised by abnormal levels of cho-1, such as Alzheimer's disease
XX
SQ Sequence 1743 BP; 407 A; 410 C; 409 G; 517 T; 0 U; 0 Other;

Query Match 78.8%; Score 1373.4; DB 4; Length 1743;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||| || ||||| ||||| ||| | ||||| || || || |||
Db      1 ATGCTTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 60

Qy     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 GTTGGAATATGGGCTGCATGGAAAACCAAAACAGCGGCAACCCAGAAGAGCACAGTGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
      ||||| ||||| || ||||| || ||||| ||||| ||||| |||||
Db    121 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC 180

Qy    181 ACCTGGGTCGGAGGAGGTTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| || ||||| ||||| ||||| ||||| |||||
Db    181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGATGGGCCAGGTTGT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db    241 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 300

Qy    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db    301 TTTTTCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCACAG 360

Qy    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTTCCTGCACTGATGGGAGAAAATGTTT 420
      ||||| ||||| || || || || || ||||| ||||| |||||
Db    361 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      ||||| ||||| || || ||||| ||||| ||||| ||||| |||||
Db    421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT 480

Qy    481 ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      || |||| || |||| ||||| ||||| ||||| || || || |||||
Db    481 GTGAACATATCGGTCAATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG 540

Qy    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600
      |||| ||||| || ||||| || ||||| || ||||| ||||| |||||
Db    541 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 600

Qy    601 ATCAGCGTCCCCTTTGCAATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
      |||| ||||| |||| ||||| ||||| ||||| ||||| |||||
Db    601 ATCAGTGTCCCCTTTGCCCTGTACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG 660

Qy    661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720
      |||| ||||| || || ||||| |||| |||| ||||| |||||
Db    661 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG 720
```

Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGGAGATTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCAAATAGATGTATTTGATGCTGTTGTCGCAAGG	1560

Qy	1561	CACAGTGAAGAAAACATGGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCTAACCCTCAGTTCAACTTTTCACCAATAAG	1680
Qy	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAAATTACAA	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

Db	727	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCCTAGTGGGTGGG	786
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Db	787	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	846
Qy	601	ATCAGCGTCCCTTTTGCAATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	847	ATCAGTGTCCCTTTTGCCCTGTCACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	906
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	907	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG	966
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	967	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	1026
Qy	781	GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	1027	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG	1086
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1087	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	1146
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1147	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	1206
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	1207	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1266
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCATGTTTGCA	1080
Db	1267	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT	1326
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1327	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1386
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1387	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1446
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Db	1447	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTACATCATCATCTTCCCACAG	1506
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1507	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1566
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380

Db	1567	TTTGGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1626
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1627	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1686
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1687	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1746
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1747	CTATTTGAAAGTGGAACCTTGCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	1806
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1807	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAT	1866
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1867	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTCACCAATAAG	1926
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1927	GAGGCCCTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAACTTACAA	1986
Qy	1741	TGA	1743
Db	1987	TGA	1989

RESULT 12

AAH49206

ID AAH49206 standard; DNA; 8760 BP.

XX

AC AAH49206;

XX

DT 26-NOV-2001 (first entry)

XX

DE Human CHOT exons 6, 7, 8 and 3' UTR region DNA.

XX

KW CHOT; human; choline transporter; chromosome 2q11-13; nootropic;

KW neuroprotective; gene therapy; antisense therapy; degenerative disease;

KW cognitive disorder; Alzheimer's disease; ds.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	1. .4853
----	-----	----------

FT		/*tag= a
----	--	----------

FT		/product= "CHOT"
----	--	------------------

FT		/note= "This sequence is interrupted by introns"
----	--	--

FT	exon	41. .194
----	------	----------

FT		/*tag= b
----	--	----------

FT		/number= 6
----	--	------------

FT	intron	195. .2456
----	--------	------------

FT		/*tag= c
----	--	----------

FT /number= 6
 FT exon 2455. .2674
 FT /*tag= d
 FT /number= 7
 FT intron 2675. .4223
 FT /*tag= e
 FT /number= 7
 FT exon 4224. .4853
 FT /*tag= f
 FT /number= 8
 FT 3'UTR 4854. .8760
 FT /*tag= g

XX
PN DE10009055-A1.

XX
PD 30-AUG-2001.

XX
PF 28-FEB-2000; 2000DE-01009055.

XX
PR 28-FEB-2000; 2000DE-01009055.

XX
PA (BRUE/) BRUESS M.
PA (BOEN/) BOENISCH H.

XX
PI Bruess M, Boenisch H;

XX
DR WPI; 2001-590709/67.

XX
PT A new gene encoding human choline transporter, designated hCHOT is
PT located on chromosome 2q11-13 and is useful to treat degenerative
PT disorders such as Alzheimer's disease.

XX
PS Disclosure; Page 9-11; 12pp; German.

XX
CC This invention describes a novel gene encoding human choline transporter,
CC designated hCHOT which is located on chromosome 2q11-13. The products of
CC the invention have nootropic and neuroprotective activity and can be used
CC for gene or antisense therapy. (I) is used to treat degenerative disease,
CC particularly cognitive disorders such as Alzheimer's disease. Sense and
CC antisense oligonucleotides derived from the gene may be used in
CC diagnostics and other techniques. This sequence represents exons 6-8 and
CC the 3' UTR fragment encoding the human CHOT protein described in the
CC invention

XX
SQ Sequence 8760 BP; 2727 A; 1619 C; 1565 G; 2849 T; 0 U; 0 Other;

Query Match 36.2%; Score 630.8; DB 5; Length 8760;
 Best Local Similarity 99.7%; Pred. No. 1.8e-172;
 Matches 632; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1110 AAATGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTGTTGTTGGAGC 1169
 | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 4220 ACAGGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTGTTGTTGGAGC 4279
 Qy 1170 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 1229
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 4280 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 4339

Qy	1230	TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC	1289
Db	4340	TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC	4399
Qy	1290	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	1349
Db	4400	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	4459
Qy	1350	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	1409
Db	4460	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	4519
Qy	1410	TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCCTTAAC	1469
Db	4520	TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCCTTAAC	4579
Qy	1470	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	1529
Db	4580	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	4639
Qy	1530	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	1589
Db	4640	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	4699
Qy	1590	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	1649
Db	4700	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	4759
Qy	1650	CATGACCCTCAGCTCAACTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	1709
Db	4760	CATGACCCTCAGCTCAACTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	4819
Qy	1710	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	1743
Db	4820	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	4853

RESULT 13

ADD50656

ID ADD50656 standard; DNA; 119040 BP.

XX

AC ADD50656;

XX

DT 15-JAN-2004 (first entry)

XX

DE BAC sequence containing hCHT gene.

XX

KW Human; high-affinity choline transporter; hCHT; chromosome 2q12;

KW cholinergic function; Parkinson's disease; Huntington's disease;

KW Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;

KW brain; cholinergic signalling; antiparkinsonian; anticonvulsant;

KW nootropic; neuroprotective; neuroleptic; bacterial artificial chromosome;

KW BAC; ds.

XX

OS Homo sapiens.

XX

XX

XX

XX

XX

PA (APPA/) APPARSUNDARAM S.

XX

XX

XX

XX

XX

XX

Best Local Similarity 99.7%; Pred. No. 7.1e-172;

Db

Qy

Db

TABLE 1

Qy	1290	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	1349
Db	30935	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	30994
Qy	1350	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	1409
Db	30995	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	31054
Qy	1410	TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAAC	1469
Db	31055	TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAAC	31114
Qy	1470	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	1529
Db	31115	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	31174
Qy	1530	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	1589
Db	31175	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	31234
Qy	1590	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	1649
Db	31235	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	31294
Qy	1650	CATGACCCTCAGCTCAACTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	1709
Db	31295	CATGACCCTCAGCTCAACTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	31354
Qy	1710	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	1743
Db	31355	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	31388

RESULT 14

ADD50651

ID ADD50651 standard; DNA; 142299 BP.

XX

AC ADD50651;

XX

DT 15-JAN-2004 (first entry)

XX

DE BAC sequence #2 containing hCHT DNA.

XX

KW Human; high-affinity choline transporter; hCHT; chromosome 2q12;

KW cholinergic function; Parkinson's disease; Huntington's disease;

KW Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;

KW brain; cholinergic signalling; antiparkinsonian; anticonvulsant;

KW nootropic; neuroprotective; neuroleptic; bacterial artificial chromosome;

KW BAC; ds.

XX

OS Homo sapiens.

XX

PN US2003114399-A1.

XX

PD 19-JUN-2003.

XX

Qy	1350	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCCTGGCTATTACCCCTGATGATAA	1409
Db	94913	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCCTGGCTATTACCCCTGATGATAA	94972
Qy	1410	TGGTATATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATCTTAAC	1469
Db	94973	TGGTATATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATCTTAAC	95032
Qy	1470	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	1529
Db	95033	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	95092
Qy	1530	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	1589
Db	95093	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	95152
Qy	1590	TCTTGTCAAAAATGAAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	1649
Db	95153	TCTTGTCAAAAATGAAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	95212
Qy	1650	CATGACCCTCAGCTCAACTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	1709
Db	95213	CATGACCCTCAGCTCAACTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	95272
Qy	1710	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	1743
Db	95273	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	95306

RESULT 15

AAF81710

ID AAF81710 standard; cDNA; 1731 BP.

XX

AC AAF81710;

XX

DT 01-JUN-2001 (first entry)

XX

DE C. elegans high affinity choline transporter protein encoding cDNA.

XX

KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
KW ss.

XX

OS Caenorhabditis elegans.

XX

FH Key Location/Qualifiers

FT CDS 1. .1731

FT /*tag= a

FT /product= "high affinity choline transporter"

XX

PN WO200116315-A1.

XX

PD 08-MAR-2001.

XX

PF 18-AUG-2000; 2000WO-JP005545.

XX

PR 27-AUG-1999; 99JP-00240642.

PR 27-DEC-1999; 99JP-00368991.

XX

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX

PI Haga T, Okuda T;

XX

DR WPI; 2001-226688/23.

DR P-PSDB; AAB74663.

XX

PT New rat and human spinal cord high affinity choline transporters, useful
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT treating Alzheimer's disease.

XX

PS Claim 3; Page 57-62; 90pp; Japanese.

XX

CC The present sequence encodes a Caenorhabditis elegans high affinity
CC choline transporter protein designated cho-1. The cho-1 protein has
CC nootropic and neuroprotective activities. The cho-1 polynucleotide and
CC protein can be used for the diagnosis of diseases related to the
CC expression of cho-1 by comparing the cho-1 polynucleotide sequence in a
CC sample to that of a control. Drug compositions containing the cho-1
CC protein or expression promoters or inhibitors of cho-1 are useful for
CC treating disorders characterised by abnormal levels of cho-1, such as
CC Alzheimer's disease

XX

SQ Sequence 1731 BP; 428 A; 373 C; 427 G; 503 T; 0 U; 0 Other;

Query Match 20.9%; Score 363.8; DB 4; Length 1731;

Best Local Similarity 55.1%; Pred. No. 5.5e-95;

Matches 862; Conservative 0; Mismatches 637; Indels 66; Gaps 5;

```
Qy      19  GGA CTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTGGTTGGAATATGGGCTGCC 78
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      16  GGTATCGTGGCCATTGTGTTCTTCTACGTGCTCATTCTTGTCTGTTGGAATATGGGCGGGT 75

Qy      79  TGGAGAACCAAAA-----ACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATC 126
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      76  AGAAAATCGAAAAGTTCAAAAAGAGCTTGAATCAGAAGCCGGCGCGGCGACGGAAGAGGTG 135

Qy     127  ATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCTACCTGG 186
      || | | | | | | | | | | | | | | | | | | | | | | | | |
Db     136  ATGTTAGCTGGGAGAAACATCGGAACCTCTTGTCGGAATTTTCACAATGACTGCCACGTGG 195

Qy     187  GTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTA 246
      || || | | | | | | | | | | | | | | | | | | | | | |
Db     196  GTTGGCGGTGCTTATATCAATGGAACCGCCGAGGCTCTGTATAATGGAGGT-----CTC 249

Qy     247  GCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTT 306
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     250  CTTGGATGTCAGGCTCCAGTTGGATATGCAATTTCCCTTGTTATGGGAGGACTACTTTTC 309

Qy     307  GCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAAATCTAT 366
      |||| | | | | | | | | | | | | | | | | | | | | | | |
Db     310  GCAAAGAAAATGCGAGAAGAAGGATATATTACAATGCTCGATCCTTTTCAGCACAAATAT 369

Qy     367  GGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAAATGTTCTGGGCT 426
      || |||| | | | | | | | | | | | | | | | | | | | | |
```

Db 370 GGCCAACGAATCGGTGGCTTGATGTATGTTCCAGCACTTCTTGGTGAAACATTCTGGACA 429
 Qy 427 GCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCAC 486
 ||||| ||| | || | | || ||| | | || | ||| |
 Db 430 GCAGCCATTCTTTCGGCACTTGGTGCAACACTGTCGGTAATTCTTGGAATCGACATGAAT 489
 Qy 487 ATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTAT 546
 || || | | || || ||||| | ||||| | || || |||||
 Db 490 GCATCAGTGACCCGTGTCGGCCTGTATTGCCGTATTCTACACATTACCCGGTGGATACTAT 549
 Qy 547 TCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTTGTAGGGCTGTGGATCAGC 606
 | || || ||||| ||||| || ||||| || || ||||| | ||
 Db 550 GCAGTCGCGTACACTGACGTGCTTCAACTATTTGCATTTTCGTCGGTTTGTGGGTTTGC 609
 Qy 607 GTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCCTGCTGTGCATGCC 666
 || || || || || || || || || || || || || || || ||
 Db 610 GTGCCGGCGGCTATGGTGCATGATGGTGCGAAGGATATTTCCAGGAATGCAG----- 661
 Qy 667 AAATACCAAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGGCTTGAT 726
 | | | | | | | | | | | | | | | | | | | | | |
 Db 662 -----GCGACTGGATTGGAGAGATTGGAGGATTCAAAGAAACATCTCTCTGGATTGAT 714
 Qy 727 AGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCTAGAGGGTTCTC 786
 | | || | | | | ||||| ||||| ||||| || || |||||
 Db 715 TGCATGCTTCTCCTTGCTTTGGAGGAATTCCATGGCAAGTGTACTTCCAAAGAGTTCTC 774
 Qy 787 TCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGCTTCCTTGGCAGCTTTTCGGGTGCCTG 846
 || || | || | || || | |||| | |||| | |||| |||||
 Db 775 TCCTCAAAAAGTCTCATGGAGCACAGACGTTGTCGTTTGTGGCGGGCGTTCGGATGCATT 834
 Qy 847 GTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAACCAG 906
 | ||||| || ||| | || || || || || || ||||| |||||
 Db 835 CTCATGGCGATTCCACCAGCGTTGATCGGTGCAATTGCCAGGAACACAGACTGGAGAATG 894
 Qy 907 ACTGCATATGGGCTTCC-----AGATCCCAAGACTACAGAAGAGGCA 948
 |||| ||| | | | | | || || | | || || || ||
 Db 895 ACTGATTATTCCCCATGGAACAATGGAACCTAAGGTGCAATCGATTCCACCGGATAAGAGA 954
 Qy 949 GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT 1008
 ||||| | | || | || ||||| || || | | || || |||||
 Db 955 AACATGGTGGTCCCGTTGGTATTCCAGTATCTTACGCCAAGATGGGTGCGCTTTATTGGA 1014
 Qy 1009 CTTGGTGCAGTTTCTGCTGCTGTTATGTGCATCAGCAGATTCTTCCATCTTGTGCAGCAAGT 1068
 || || ||||| || ||||| ||||| ||||| || || | |||||
 Db 1015 CTCGGCGCAGTGTCGGCTGCTGTAATGTGCATCTGCAGATTCTGTACTATCAGCAGCA 1074
 Qy 1069 TCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAA 1128
 || ||||| || ||||| |||| | | | | || || || |||||
 Db 1075 TCAATGTTTGTCTACAACATCTGGAAGCTCACAATTGCCCCACGCGTCTGAAAAAGAA 1134
 Qy 1129 ATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTGGAGCATCTGCAACAGCCATGGCC 1188
 | | | || |||| | | | | |||| | |||| | |||||
 Db 1135 GTGATAAATTGTGATGAGAATAGCCATCATCTGTGTTGGTATCATGGCAACCATCATGGCA 1194
 Qy 1189 TTGCTGACGAAAAGTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTT 1248
 | | | || | ||||| ||||| || || | || | || || || ||
 Db 1195 CTTACCATTCAATCCATCTATGGGCTTTGGTATCTTGTGCAGATTTGGTCTACGTCATA 1254

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2004, 20:43:21 ; Search time 151 Seconds
(without alignments)
6405.824 Million cell updates/sec

Title: US-10-069-541-5
Perfect score: 1743
Sequence: 1 atggctttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					Description
Result	Query						
No.	Score	Match	Length	DB	ID		

1	1738.2	99.7	1743	4	US-09-657-252-1		Sequence 1, Appli
2	47.6	2.7	7218	1	US-08-232-463-14		Sequence 14, Appl
3	41	2.4	1857	4	US-09-640-198D-3		Sequence 3, Appli
4	41	2.4	2839	4	US-08-595-553A-1		Sequence 1, Appli
5	39.6	2.3	474	4	US-09-621-976-18033		Sequence 18033, A
6	39.2	2.2	1506	4	US-09-328-352-2245		Sequence 2245, Ap
7	39	2.2	558	4	US-09-328-352-3451		Sequence 3451, Ap
8	38.2	2.2	2028	4	US-10-162-012-28		Sequence 28, Appl
9	38.2	2.2	2326	4	US-10-162-012-26		Sequence 26, Appl
c 10	38.2	2.2	1830121	4	US-09-557-884-1		Sequence 1, Appli
c 11	38.2	2.2	1830121	4	US-09-643-990A-1		Sequence 1, Appli

	12	38	2.2	1932	4	US-09-640-198D-1	Sequence 1, Appli
	13	37.4	2.1	4344	4	US-09-601-198-165	Sequence 165, App
c	14	36.6	2.1	4160	4	US-09-134-218-1	Sequence 1, Appli
	15	36.6	2.1	148567	4	US-09-801-876B-3	Sequence 3, Appli
	16	36.6	2.1	148567	4	US-10-254-869-3	Sequence 3, Appli
	17	35.6	2.0	447	4	US-09-621-976-12063	Sequence 12063, A
c	18	35.6	2.0	2397	4	US-09-221-017B-272	Sequence 272, App
	19	35.2	2.0	1461	4	US-09-543-681A-2066	Sequence 2066, Ap
	20	35.2	2.0	2238	1	US-07-841-651-1	Sequence 1, Appli
	21	34.8	2.0	902	4	US-09-671-317-37	Sequence 37, Appl
	22	34.8	2.0	1593	4	US-09-134-001C-1673	Sequence 1673, Ap
c	23	34.8	2.0	12482	4	US-09-512-563C-25	Sequence 25, Appl
c	24	34.8	2.0	25002	4	US-08-961-527-48	Sequence 48, Appl
	25	34.8	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	26	34.6	2.0	561	4	US-09-107-532A-3215	Sequence 3215, Ap
	27	34.6	2.0	1005	4	US-09-107-532A-3570	Sequence 3570, Ap
	28	34.6	2.0	2847	4	US-09-484-970B-22	Sequence 22, Appl
c	29	34.6	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	30	34.2	2.0	1515	4	US-09-071-035-431	Sequence 431, App
	31	34.2	2.0	1803	4	US-09-071-035-429	Sequence 429, App
c	32	34	2.0	1109	4	US-08-956-171E-222	Sequence 222, App
c	33	34	2.0	392000	4	US-10-027-983-11	Sequence 11, Appl
c	34	33.8	1.9	369	4	US-09-543-681A-628	Sequence 628, App
c	35	33.8	1.9	3172	1	US-07-741-940-3	Sequence 3, Appli
c	36	33.8	1.9	3172	1	US-08-289-548A-3	Sequence 3, Appli
c	37	33.8	1.9	3172	1	US-08-452-654-3	Sequence 3, Appli
c	38	33.8	1.9	3172	1	US-08-452-655B-3	Sequence 3, Appli
c	39	33.8	1.9	3172	3	US-08-450-582-3	Sequence 3, Appli
c	40	33.8	1.9	3172	4	US-08-449-731-3	Sequence 3, Appli
	41	33.8	1.9	176373	3	US-09-128-155-17	Sequence 17, Appl
	42	33.6	1.9	84495	4	US-09-797-906-3	Sequence 3, Appli
	43	33.4	1.9	1626	4	US-09-328-352-602	Sequence 602, App
	44	33.4	1.9	3593	4	US-09-404-627-3	Sequence 3, Appli
	45	33.4	1.9	4205	4	US-09-404-627-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-657-252-1

; Sequence 1, Application US/09657252

; Patent No. 6500643

; GENERAL INFORMATION:

; APPLICANT: Wu, Dong-Hai

; APPLICANT: Gu, Yunrong

; APPLICANT: Millard, William

; APPLICANT: He, Yun-Je

; TITLE OF INVENTION: Human High Affinity Choline Transporter cDNA

; FILE REFERENCE: MBHB00-639

; CURRENT APPLICATION NUMBER: US/09/657,252

; CURRENT FILING DATE: 2000-09-07

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 1743

; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-09-657-252-1

Query Match 99.7%; Score 1738.2; DB 4; Length 1743;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180

Qy    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300

Qy    301 TTCTTTGCAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTCTTTGCAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360

Qy    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480

Qy    481 ATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540

Qy    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600

Qy    601 ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660

Qy    661 CATGCCAAATACCAAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CATGCCAAATACCAAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG 720
```

Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560

QY 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
 |||
 Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
 QY 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
 |||
 Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
 QY 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 |||
 Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAA 1740
 QY 1741 TGA 1743
 |||
 Db 1741 TGA 1743

RESULT 2

US-08-232-463-14

; Sequence 14, Application US/08232463

; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/232,463

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,313

; FILING DATE:

; APPLICATION NUMBER: EP 91 114 300.6

; FILING DATE: 26-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 30472/114 IMMU

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)836-9300

; TELEFAX: (703)683-4109

; TELEX: 899149

; INFORMATION FOR SEQ ID NO: 14:

```
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 7218 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   CLONE: pTZgpt-Fls
US-08-232-463-14
```

```
Query Match          2.7%; Score 47.6; DB 1; Length 7218;
Best Local Similarity 6.0%; Pred. No. 0.00054;
Matches 23; Conservative 200; Mismatches 159; Indels 0; Gaps 0;
```

```
Qy      692 CTGTTGACTCATCTGAAGTCTACTCTTGGCTTGATAGTTTCTGTTGTTGATGCTGGGTG 751
          :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      1088 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1147

Qy      752 GAATCCCATGGCAAGCATACTTTTCAGAGGGTTCTCTCTTCTCCTCAGCCACCTATGCTC 811
          :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      1148 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1207

Qy      812 AAGTGCTGTCCTTCCTGGCAGCTTTCGGGTGCCTGGTGATGGCCATCCAGCCATACTCA 871
          : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      1208 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1267

Qy      872 TTGGGGCCATTGGAGCATCAACAGACTGGAACCAGACTGCATATGGGCTTCCAGATCCCA 931
          :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      1268 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1327

Qy      932 AGACTACAGAAGAGGCAGACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGT 991
          :: : : : : :: :: :: :: :: :: :: :: :: :: :: ::
Db      1328 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1387

Qy      992 ATATTTCTTTCTTTGGTCTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTT 1051
          : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      1388 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCAAATTCTT 1447

Qy      1052 CCATCTTGTCAGCAAGTTCCAT 1073
          | ||| | | | | || |||
Db      1448 CTATCTCTTTAACTACTTGCAT 1469
```

RESULT 3

US-09-640-198D-3

; Sequence 3, Application US/09640198D

; Patent No. 6586411

; GENERAL INFORMATION:

; APPLICANT: Russell, Stephen

; APPLICANT: Kay Whye, Peng

; TITLE OF INVENTION: System for Monitoring the Location of

; TITLE OF INVENTION: Transgenes

; FILE REFERENCE: 07039-295001

; CURRENT APPLICATION NUMBER: US/09/640,198D

; CURRENT FILING DATE: 2000-08-16

; PRIOR APPLICATION NUMBER: US 60/149,168

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-640-198D-3

Query Match 2.4%; Score 41; DB 4; Length 1857;
Best Local Similarity 49.8%; Pred. No. 0.026;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```
Qy      404 TGATGGGAGAAATGTTCTGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCG 463
      || |||      ||| | | | | | | | | | | | | | | |
Db      413 TGGTGGCCACGATGCTGTATACAGGCATCGTGATCTACGCGCCTGCGCTCATCCTGAACC 472

Qy      464 TGATCATCGATGTGGATATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGT 523
      | | ||      |||| | | | | | | | | | | | | | | |
Db      473 AAGTGACCGGGTTGGACATCTGGGCATCGCTCCTGTCCACAGGAATCATCTGCACCTTGT 532

Qy      524 ACACACTGGTGGGAGGGCTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGA 583
      ||||      |||| | | | | | | | | | | | | | | |
Db      533 ACACTACCGTGGGTGGTATGAAGGCCGTGGTCTGGACAGATGTGTTCCAGGTTGTGGTAA 592

Qy      584 TTTTGTAGGGCTGTGGATCAGCGTCCCC 612
      | | || | | | | | | | | | | | | | |
Db      593 TGCTCGTTGGCTTCTGGGTGATCCTGGCC 621
```

RESULT 4

US-08-595-553A-1

; Sequence 1, Application US/08595553A

; Patent No. 6391579

; GENERAL INFORMATION:

; APPLICANT: NANCY CARRASCO, ET AL.

; TITLE OF INVENTION: THYROID SODIUM/IODIDE SYMPORTER AND

; TITLE OF INVENTION: NUCLEIC ACID ENCODING SAME

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN

; STREET: 90 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: U.S.A.

; ZIP: 10016

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE

; MEDIUM TYPE: DISKETTE

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/595,553A

; FILING DATE: FEBRUARY 1, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: CRAIG J. ARNOLD

; REGISTRATION NUMBER: 34,287

; REFERENCE/DOCKET NUMBER: 96700/393

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2839
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: OLIGONUCLEOTIDE
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: RAT
; INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER
US-08-595-553A-1

```

```

Query Match          2.4%; Score 41; DB 4; Length 2839;
Best Local Similarity 49.8%; Pred. No. 0.035;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```

```

Qy      404 TGATGGGAGAAATGTTCTGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCG 463
        || |||      ||| | | | | | | | | | | | | | | | |
Db      522 TGGTGGCCACGATGCTGTATACAGGCATCGTGATCTACGCGCCTGCGCTCATCCTGAACC 581

Qy      464 TGATCATCGATGTGGATATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGT 523
        | | | | | ||| | | | | | | | | | | | | | | |
Db      582 AAGTGACCGGGTTGGACATCTGGGCATCGCTCCTGTCCACAGGAATCATCTGCACCTTGT 641

Qy      524 ACACACTGGTGGGAGGGCTCTATTCTGTGGCCTACACTGATGTTCGTTTCAGCTCTTTTGCA 583
        ||||      |||| | | | | | |||| | | |||| | | | |
Db      642 ACACTACCGTGGGTGGTATGAAGCCGTGGTCTGGACAGATGTGTTCCAGGTTGTGGTAA 701

Qy      584 TTTTGTAGGGCTGTGGATCAGCGTCCCC 612
        | | | | | | | | | | | | | |
Db      702 TGCTCGTTGGCTTCTGGGTGATCCTGGCC 730

```

RESULT 5

```

US-09-621-976-18033
; Sequence 18033, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18033
; LENGTH: 474
; TYPE: DNA

```


Query Match 2.3%; Score 39.6; DB 4; Length 474;
Best Local Similarity 13.4%; Pred. No. 0.03;
Matches 42; Conservative 134; Mismatches 138; Indels 0; Gaps 0;

```

RESULT 6
US-09-328-352-2245
; Sequence 2245, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2245
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2245

```

Query Match 2.2%; Score 39.2; DB 4; Length 1506;
Best Local Similarity 51.1%; Pred. No. 0.085;

Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

Qy 1008 TCTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTTCAGCAAG 1067
    ||| | ||| ||| ||||| ||||| || | | | ||| |
Db 1005 TCTAGCTGCTATTTTAGCTGCGGTATGAGTACATTAAGCTGTCAGCTTTTGGTATGTTC 1064

Qy 1068 TTCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGA 1127
    | | | | | | | | | | | | | | | | | | | |
Db 1065 AAGTGCCTAACTGAAGATTGTACAAAGGCTTCATTCGTAAAAATGCATCTCAAAAAGA 1124

Qy 1128 AATCGTTTGGGTTATGCGAATCACAGTGTGTGTTGGAGCATCTGCAACAGCCATGGC 1187
    | || ||||| ||| ||| ||| ||| ||| | |||| | |||
Db 1125 GCTTGTATGGGTTGGGCGTATCATGGTGCCTGCAATTGCCGTTCTAGCAATTGTGCTTGC 1184

```

RESULT 7

US-09-328-352-3451

; Sequence 3451, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 3451

; LENGTH: 558

; TYPE: DNA

; ORGANISM: Acinetobacter baumannii

US-09-328-352-3451

Query Match 2.2%; Score 39; DB 4; Length 558;

Best Local Similarity 50.3%; Pred. No. 0.051;

Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```

Qy 1455 TACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGG 1514
    || |||| | | ||| | ||| | ||| ||| ||| |
Db 341 TAAATCAAATGATGCAAATGCTTCATGGCTGATGTTGCAAACCTTCAACTTTTCAAGATG 400

Qy 1515 AACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAA 1574
    | | | | | | | | | | | | | | | | | | | |
Db 401 GCCGTAGTCATCTGAATGCGGCAAAGCTCAAGGTGAAGTTTCAGAAGCAAGCAGATGGAA 460

Qy 1575 CATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGATGAAGTTGCACTTGT 1634
    ||||| || || || || | ||| ||| ||| | | ||| | |
Db 461 CATGGAAAATTAAACATTTCCAAACACAGAATATTTTCAGTCGTCGGTATCGCATTGGC 520

Qy 1635 GAAGCCACGAC 1645
    ||| | | |
Db 521 AAAGTGAAGCC 531

```

RESULT 8

US-10-162-012-28

; Sequence 28, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10/162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-162-012-28

Query Match 2.2%; Score 38.2; DB 4; Length 2028;
 Best Local Similarity 46.1%; Pred. No. 0.22;
 Matches 239; Conservative 0; Mismatches 273; Indels 7; Gaps 3;

```

Qy      79 TGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATCATAGTTGGTGGC 138
      |||| | ||||| | | | | ||| ||| | | ||| |
Db     126 TGGACTATGGTCCACAGTGAAGACCAAAAGAGACACAGTGAAAGGCTACTTCCTGGCTGA 185

Qy     139 CGAGATATTGGTTTATTGGTTGGTGGATTTA-CCATGACAGCTACCTGGGTGCGGAGGAGG 197
      | || |||| | ||| |||| | | ||| | ||| | ||| | |||
Db     186 AGGGAACATGGTGTGGTGGCCAGTGGGTGCATCCTTGTTTGCCAGCAATGTTGGAAGTGG 245

Qy     198 GTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTAGCTTGGGCTCA 257
      || ||| |||| | | | ||| | | | | | | | | | ||| |
Db     246 ACATTTTCATTGGCCTGGCAGGGTCAGGTGCTGCTACGGGCATTTCTGTA---TCAGCTTA 302

Qy     258 GGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGCAAAACCTAT 317
      | || | ||| | | | | |||| | | | | | | | | | ||
Db     303 TGAACCTTAATGGCTTGTTTTCTGTGCTGATGTTGGCCTGGATCTTCCTACCCATCTACAT 362

Qy     318 GCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGGAAAACGCAT 377
      | ||| || | | | ||| ||| ||| ||| ||| | |||
Db     363 TGCTGGTCAGGTCACCACGATGCCAGAATACCTACGGAAGCGCTTCGGTGGCATCAGAAT 422

Qy     378 GGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTCTGGGCTGCAGCAATTTT 437
      | |||| | | | || | | | ||| ||| | | |
Db     423 CCCCATCATCCTGGCTGTACTCTACCTATTTATCTACATCTTCACCAAGATCTCGGTAGA 482

Qy     438 CTCTGCTTTGGGAGCCACCATCAGCGTGATCATCG---ATGTGGATATGCACATTTCTGT 494
      | | | |||| | ||| | | | | |||| | | | | | |
Db     483 CATGTATGCAGGTGCCATCTTCATCCAGCAGTCTTCGCACCTGGATCTGTACCTGGCCAT 542

Qy     495 CATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTCTGTGGC 554
      | ||| | | | |||| | | ||| ||| | ||||
Db     543 AGTTGGGCTACTGGCCATCACTGCTGTATACACGGTTGCTGGTGGCCTGGCTGCTGTGAT 602

Qy     555 CTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGG 593
      ||||| |||| | | ||| | ||| ||| ||||
Db     603 CTACACGGATGCCCTGCAGACGCTGATCATGCTTATAGG 641
  
```

RESULT 9

US-10-162-012-26

; Sequence 26, Application US/10162012

; Patent No. 6682597

; GENERAL INFORMATION:

; APPLICANT: Curtis, Rory A.J.

; APPLICANT: Silos-Santiago, Inmaculada

; APPLICANT: Gu, Wei

; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS

; FILE REFERENCE: 10448-190001

; CURRENT APPLICATION NUMBER: US/10/162,012

; CURRENT FILING DATE: 2002-06-04

; PRIOR APPLICATION NUMBER: US 60/209,845

; PRIOR FILING DATE: 2000-06-06

; PRIOR APPLICATION NUMBER: US 09/875,321

```

; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 2326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (178)...(2202)
US-10-162-012-26

```

```

Query Match          2.2%; Score 38.2; DB 4; Length 2326;
Best Local Similarity 46.1%; Pred. No. 0.24;
Matches 239; Conservative 0; Mismatches 273; Indels 7; Gaps 3;

```

```

Qy      79 TGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATCATAGTTGGTGGC 138
        |||| |      |||||      | | |      | ||| |||      | | ||| |
Db      303 TGGACTATGGTCCACAGTGAAGACCAAAAGAGACACAGTGAAAGGCTACTTCCTGGCTGA 362

Qy      139 CGAGATATTGGTTTATTGGTTGGTGGATTTA-CCATGACAGCTACCTGGGTCGGAGGAGG 197

```

Db	363	AGGGAACATGGTGTGGTGGCCAGTGGGTGCATCCTTGTTTGCCAGCAATGTTGGAAGTGG	422
Qy	198	GTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTAGCTTGGGCTCA	257
Db	423	ACATTTTCATTGGCCTGGCAGGTCAGGTGCTGCTACGGGCATTTCTGTA---TCAGCTTA	479
Qy	258	GGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGCAAACCTAT	317
Db	480	TGAACCTTAATGGCTTGTTTTCTGTGCTGATGTTGGCCTGGATCTTCCTACCCATCTACAT	539
Qy	318	GCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGGAAAACGCAT	377
Db	540	TGCTGGTCAGGTCACCACGATGCCAGAATACCTACGGAAGCGCTTCGGTGGCATCAGAAT	599
Qy	378	GGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTCTGGGCTGCAGCAATTTT	437
Db	600	CCCCATCATCCTGGCTGTACTCTACCTATTTATCTACATCTTCACCAAGATCTCGGTAGA	659
Qy	438	CTCTGCTTTGGGAGCCACCATCAGCGTGATCATCG---ATGTGGATATGCACATTTCTGT	494
Db	660	CATGTATGCAGGTGCCATCTTCATCCAGCAGTCTTCGCACCTGGATCTGTACCTGGCCAT	719
Qy	495	CATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTCTGTGGC	554
Db	720	AGTTGGGCTACTGGCCATCACTGCTGTATACAGGTTGCTGGTGGCCTGGCTGCTGTGAT	779
Qy	555	CTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGG	593
Db	780	CTACACGGATGCCCTGCAGACGCTGATCATGCTTATAGG	818

RESULT 10

US-09-557-884-1/c

; Sequence 1, Application US/09557884

; Patent No. 6506581

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3 1/2 inch diskette

; COMPUTER: Dell Pentium

; OPERATING SYSTEM: MS DOS v6.22

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/557,884

; FILING DATE: 25-Apr-2000

Owen White

Qy	1129	ATCGTTTGGGTATGCGAATCACAGTGT	TTTGTGTTGGAGCATCTGCAACAGCCATGGC-	1187
Db	1428665	CTCGTATGGCTTGGCAGAATTATGGTGT	TATGTTAGTTATTGCCGCACTTGCTATCTGGATCGCA	
	1428606			
Qy	1188	CTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGT		1247
Db	1428605	CAAGATGAAAACAGCAAAGTATTAAACTTGTAGAATTGCTTGGGCGGGGTTTGGTAGT		
	1428546			
Qy	1248	TATCTTCCCCCAGCTGCTTTTGTGTACTCTTT		1278
Db	1428545	GCATTTGGCCCTGTTGTACTTTTCTCTCTTT		1428515

RESULT 12

US-09-640-198D-1

; Sequence 1, Application US/09640198D

; Patent No. 6586411

; GENERAL INFORMATION:

; APPLICANT: Russell, Stephen

; APPLICANT: Kay Whye, Peng

; TITLE OF INVENTION: System for Monitoring the Location of

; TITLE OF INVENTION: Transgenes

; FILE REFERENCE: 07039-295001

: CURRENT APPLICATION NUMBER: US/09/640,198D

; CURRENT FILING DATE: 2000-08-16

PRIOR APPLICATION NUMBER: US 60/149,168

; PRIOR FILING DATE: 1999-08-17

```

; NUMBER OF SEQ ID NOS: 34

```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```

; SEQ ID NO 1

```

```

; LENGTH: 1932

```

; TYPE: DNA

; ORGANISM: Homo Sapiens

US-09-640-198D-1

Query Match 2.2%; Score 38; DB 4; Length 1932;

Best Local Similarity 51.8%; Pred. No. 0.24;

Best Local Similarity 51.8%; Fied. NO. 0.24;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY	437	TCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCACATTTCTGTCA	496
Db	446	TCTACGCACCGGCCCTCATCCTGAACCAAGTGACCGGGCTGGACATCTGGGCGTCGCTCC	505
QY	497	TCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTCTGTGGCCT	556
Db	506	TGTCCACCGGAATTATCTGCACCTTCTACACGGCTGTGGGCGGCATGAAGGCTGTGGTCT	565
QY	557	ACACTGATGTCGTTACAGCTCTTTTGCATTTTGTAGGGCTGTGGAT	602
Db	566	GGACTGATGTGTTCCAGGTCTGGTGATGCTAAGTGCGTCTGGGT	611

RESULT 13

US-09-601-198-165

; Sequence 165, Application US/09601198

```
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellson Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 165
; LENGTH: 4344
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-165
```

```
Query Match          2.1%; Score 37.4; DB 4; Length 4344;
Best Local Similarity 48.0%; Pred. No. 0.64;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
```

```
Qy      1401 TGATGATAATGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATC 1460
          || |||| |      || || |   || |   || || || |   |   |   |
Db      864 TGTGATCAATCAGTAGATTTTTTAAAGTAAATATTGAAGCATTAAATTAATCATCAACC 923

Qy      1461 ATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTT 1520
          | |   || || |   || || |   || |   || || || |   || || |
Db      924 ACTTAAAAACACAACATGAAACGATTTTATTAACAAAAATGTTACAGATATTAGTGCTTT 983

Qy      1521 GCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGA 1580
          || || || |   || || || || || |   || || || || || || || ||
Db      984 AAGTAACCTATTAGAAATTTTTGAACTAATGAAATTACAAATAATGAATGAAACCAATT 1043

Qy      1581 TAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGATGAA 1623
          |   || || |   | || || || |   || || |   |   || ||
Db      1044 AATTACGATTTTAATTAATCATGCACCTATTGATAAAATTGAA 1086
```

RESULT 14

US-09-134-218-1/c

```
; Sequence 1, Application US/09134218A
; Patent No. 6312926
; GENERAL INFORMATION:
; APPLICANT: Shatkin, Aaron J.
; APPLICANT: Pillutla, Renuka
; APPLICANT: Reinberg, Danny
; APPLICANT: Yu, Zheng
; APPLICANT: Moldanado, Edio
; TITLE OF INVENTION: mRNA CAPPING ENZYMES AND USES THEREOF
; FILE REFERENCE: 601-1-079 ss
; CURRENT APPLICATION NUMBER: US/09/134,218A
```

; CURRENT FILING DATE: 1998-08-14
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 4160
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-09-134-218-1

Query Match 2.1%; Score 36.6; DB 4; Length 4160;
 Best Local Similarity 46.6%; Pred. No. 1.1;
 Matches 117; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```

Qy      1417 TATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAACCAACATT 1476
          || || || || || || || || || || || || || || || || || || || ||
Db      2323 TAAACTTAAATCCAACATTTAAAAAACTCAATATGCTTACAGCTTCAGATTGCTAGTTA 2264

Qy      1477 TGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGAT 1536
          || || || || || || || || || || || || || || || || || || || ||
Db      2263 TGAATCAAATGTAAAGGTATCTATTCACATACAAACAGGCTCTATTTCACTTAACCTTCAAT 2204

Qy      1537 GTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAATTCTTGTC 1596
          | || | || | || | || | || | || | || | || | || | || | ||
Db      2203 CTGATTTAACCTTTGGGTATTTCAATCTGTAGACTCCACAGGTAAGGCTGAATTTATTC 2144

Qy      1597 AAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAGCATGACC 1656
          | || || || || || || || || || || || || || || || || || || ||
Db      2143 AGGTATAAATAAAATATTTAGGTCCATGATGTACTGTAGTTCCAAGGAAACCAATGTAC 2084

Qy      1657 CTCAGCTCAAC 1667
          | | | ||
Db      2083 CAAATATATAC 2073
  
```

RESULT 15

US-09-801-876B-3

; Sequence 3, Application US/09801876B
 ; Patent No. 6492155
 ; GENERAL INFORMATION:
 ; APPLICANT: YE, Jane et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001160
 ; CURRENT APPLICATION NUMBER: US/09/801,876B
 ; CURRENT FILING DATE: 2001-03-09
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 148567
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(148567)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-801-876B-3

Query Match 2.1%; Score 36.6; DB 4; Length 148567;
 Best Local Similarity 46.1%; Pred. No. 12;
 Matches 123; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

```

Qy      1416 ATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATTCTTAACCAACAT 1475
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31673 AGATAATCAGTTGTTTTAACTTTTAATTTAAGCAGTAGCAGAATGACTTTTTGGGAACCTT 31732

Qy      1476 TTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGA 1535
          | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31733 AGGAATTTGGAAACCTTTTTATTCTATGTATTGAATATCAACTATGTAATTTAGTCTAAG 31792

Qy      1536 TGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAATTCTTGT 1595
          | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31793 GTTATATGCTAGAAACATTTCAAAAACGAAAGCAGCAGCAATGACATCAAAAATGCATGT 31852

Qy      1596 CAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAGCATGAC 1655
          | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31853 CAAAAGCAAATGGTTTTAAATAGAAATACATCATTTTAAACAATCTTGAAGTTTAAAGAT 31912

Qy      1656 CCTCAGCTCAACTTTTACCAATAAAGA 1682
          | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31913 CCTATAAAAATCACAAACCCAGAAGGA 31939
  
```

Search completed: September 29, 2004, 02:05:54
 Job time : 159 secs

OM nucleic - nucleic search, using sw model

Run on: September 29, 2004, 00:10:08 ; Search time 1188 Seconds
(without alignments)
7431.150 Million cell updates/sec

Title: US-10-069-541-5
Perfect score: 1743
Sequence: 1 atggcgtttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result 8
Query

No.	Score	Match	Length	DB	ID	Description
1	1743	100.0	1743	10	US-09-911-077A-1	Sequence 1, Appli
2	1743	100.0	1813	10	US-09-911-077A-9	Sequence 9, Appli
3	1394.2	80.0	4904	10	US-09-911-077A-5	Sequence 5, Appli
4	1375	78.9	1743	10	US-09-911-077A-3	Sequence 3, Appli
5	1375	78.9	1743	10	US-09-911-077A-23	Sequence 23, Appl
6	630.8	36.2	119040	10	US-09-911-077A-19	Sequence 19, Appl
7	630.8	36.2	142299	10	US-09-911-077A-14	Sequence 14, Appl
8	376.6	21.6	1833	13	US-10-241-784-1	Sequence 1, Appli
9	363.8	20.9	1985	10	US-09-911-077A-7	Sequence 7, Appli
10	242.6	13.9	1461	9	US-09-974-300-501	Sequence 501, App
c 11	180.8	10.4	119040	10	US-09-911-077A-19	Sequence 19, Appl
c 12	180.8	10.4	142299	10	US-09-911-077A-14	Sequence 14, Appl
13	155	8.9	455	9	US-09-864-761-1838	Sequence 1838, Ap
c 14	118.6	6.8	943	13	US-10-027-632-120553	Sequence 120553,
c 15	118.6	6.8	943	16	US-10-027-632-120553	Sequence 120553,
16	72	4.1	96	9	US-09-864-761-18589	Sequence 18589, A
17	60	3.4	60	10	US-09-908-975-10249	Sequence 10249, A
18	53.8	3.1	65	10	US-09-908-975-26842	Sequence 26842, A
19	41.4	2.4	61396	17	US-10-322-281-483	Sequence 483, App
20	41	2.4	1857	16	US-10-428-868-3	Sequence 3, Appli
21	41	2.4	2839	9	US-09-995-007-1	Sequence 1, Appli
c 22	39.8	2.3	666	13	US-10-027-632-137101	Sequence 137101,
c 23	39.8	2.3	666	16	US-10-027-632-137101	Sequence 137101,
24	39.6	2.3	2028	9	US-09-733-630-1	Sequence 1, Appli
25	39.6	2.3	2456	9	US-09-733-630-3	Sequence 3, Appli
26	39.6	2.3	2475	13	US-10-302-172-932	Sequence 932, App
c 27	39.4	2.3	578	13	US-10-027-632-192644	Sequence 192644,
c 28	39.4	2.3	578	16	US-10-027-632-192644	Sequence 192644,
29	39.2	2.2	1173	17	US-10-437-963-19138	Sequence 19138, A
c 30	39.2	2.2	1254	17	US-10-437-963-19139	Sequence 19139, A
31	38.2	2.2	2028	9	US-09-928-530-3	Sequence 3, Appli
32	38.2	2.2	2028	15	US-10-162-012-28	Sequence 28, Appl
33	38.2	2.2	2028	16	US-10-162-102-28	Sequence 28, Appl
34	38.2	2.2	2326	9	US-09-928-530-1	Sequence 1, Appli
35	38.2	2.2	2326	15	US-10-162-012-26	Sequence 26, Appl
36	38.2	2.2	2326	16	US-10-162-102-26	Sequence 26, Appl
c 37	38.2	2.2	1830121	15	US-10-329-960-1	Sequence 1, Appli
c 38	38.2	2.2	1830121	16	US-10-329-670-1	Sequence 1, Appli
39	38	2.2	650	13	US-10-027-632-190544	Sequence 190544,
40	38	2.2	650	16	US-10-027-632-190544	Sequence 190544,
41	38	2.2	1932	16	US-10-428-868-1	Sequence 1, Appli
c 42	38	2.2	4936	12	US-09-809-665A-29	Sequence 29, Appl
c 43	37.8	2.2	6306	15	US-10-239-676-129	Sequence 129, App
c 44	37.6	2.2	867	13	US-10-142-426-20	Sequence 20, Appl
c 45	37.6	2.2	867	15	US-10-123-155-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-911-077A-1

; Sequence 1, Application US/09911077A

; Publication No. US20030114399A1

; GENERAL INFORMATION:

```
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARSUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-09-911-077A-1
```

```
Query Match          100.0%; Score 1743; DB 10; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
        |||
Db      1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
        |||
Db     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
        |||
Db    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180

Qy    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
        |||
Db    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
        |||
Db    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300

Qy    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
        |||
Db    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360

Qy    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420
        |||
Db    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
        |||
Db    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480

Qy    481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
        |||
Db    481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
```

Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG 600
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG 600

Qy 601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTG 660
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTG 660

Qy 661 CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 720
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 661 CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 720

Qy 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG 780
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG 780

Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTTCGGG 840
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTTCGGG 840

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGTTCCATGTTTGCA 1080
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGTTCCATGTTTGCA 1080

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140

Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG 1260
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG 1260

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440
 |||
 Db 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440

Qy 1441 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500
 |||
 Db 1441 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560
 |||
 Db 1501 CTATTTGAAAGTGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
 |||
 Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
 |||
 Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
 |||
 Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740

Qy 1741 TGA 1743
 |||
 Db 1741 TGA 1743

RESULT 2

US-09-911-077A-9

; Sequence 9, Application US/09911077A
 ; Publication No. US20030114399A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BLAKELY, RANDY D.
 ; APPLICANT: APPARSUNDARAM, SUBRAMANIAM
 ; APPLICANT: FERGUSON, SHAWN
 ; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
 ; FILE REFERENCE: VBLT:008US
 ; CURRENT APPLICATION NUMBER: US/09/911,077A
 ; CURRENT FILING DATE: 2001-07-23
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 1813
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (19)..(1761)
 US-09-911-077A-9

Query Match 100.0%; Score 1743; DB 10; Length 1813;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Db	19		ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	78
Qy	61		GTTGGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	79		GTTGGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	138
Qy	121		GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	139		GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	198
Qy	181		ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	199		ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	258
Qy	241		GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	259		GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	318
Qy	301		TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	319		TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	378
Qy	361		ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Db	379		ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	438
Qy	421		TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	439		TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	498
Qy	481		ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	499		ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	558
Qy	541		CTCTATTCTGTGGCCTACACTGATGTGCTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	559		CTCTATTCTGTGGCCTACACTGATGTGCTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	618
Qy	601		ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	619		ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	678
Qy	661		CATGCCAAATACCAAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	679		CATGCCAAATACCAAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	738
Qy	721		CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	739		CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	798
Qy	781		GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	799		GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	858
Qy	841		TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900

Db	859	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGGCCATTGGAGCATCAACAGACTGG	918
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	919	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	978
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	979	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1038
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1039	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1098
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1099	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1158
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1159	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1218
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1219	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1278
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1279	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1338
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1339	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1398
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1399	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1458
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1459	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1518
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1519	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1578
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1579	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1638
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1639	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1698
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1699	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1758

Qy	1741	TGA	1743
Db	1759	TGA	1761

RESULT 3

US-09-911-077A-5

; Sequence 5, Application US/09911077A

; Publication No. US20030114399A1

; GENERAL INFORMATION:

APPLICANT: BLAKELY, RANDY D.

; APPLICANT: APPARSUNDARAM, SUBRAMANIAM

; APPLICANT: FERGUSON, SHAWN

; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA

FILE REFERENCE: VBLT:008US

FILE REFERENCE: VDEF:000000
CURRENT APPLICATION NUMBER: US/09/911,077A

; CURRENT FILING DATE: 2001-07-23

```

; CURRENT FILING DATE: 11/15/2011
; NUMBER OF SEQ ID NOS: 27

```

```
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
```

```

; SOFTWARE.
; SEQ ID NO 5

```

```
; SEQ ID NO 3
; LENGTH: 4904
```

```

; LENGTH: 43
; TYPE: DNA

```

```
; TYPE: DNA
; ORGANISM: Rattus norvegicus
```

```

; ORGANISM
: FEATURE:

```

```
; FEATURE:
: NAME/KEY: CDS
```

```

; NAME/KEY: CDS
; LOCATION: (224)..(1966)

```

US-09-911-077A-5

Query Match 80.0%; Score 1394.2; DB 10; Length 4904;
 87.5%; Prod. No. 0;

Query Match 89.5%;
Best Local Similarity 87.5%; Pred. No. 0;

Best Local Similarity 87.5%; Pred. NO. 0;
Matches 1525; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTCGCTG	60
Db	224	ATGCCTTTCCATGTAGAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTTATATTTCTG	283
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACC AAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	284	GTTGGAATATGGGCTGCATGGAAAACC AAAAACAGCGGTAATGCAGAAGAACGCAGCGAA	343
Qy	121	GCCATCATAGTTGGTG G C C G A G A T A T T G G T T T A T T G G T T G G T G G A T T T A C C A T G A C A G C T	180
Db	344	GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTTGGTTGGTGGTTTTACCATGACAGCC	403
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	404	ACCTGGGTTGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT	463
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	464	GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG	523
Qy	301	TTCTTTGCAA AACCTATGCGTTCAAAGGGGATATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	524	TTTTTTGCAA AACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG	583

QY	361	ATCTATGGAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	584	ATCTATGGAAAGCGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTT	643
QY	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	644	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT	703
QY	481	ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	704	GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCTCGTGGGAGGG	763
QY	541	CTCTATTCTGTGGCCTACACTGATGTCTGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	764	CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG	823
QY	601	ATCAGCGTCCCCTTTGCATTGTCTCACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTG	660
Db	824	ATCAGTGTCCCATTTGCCCTGTCTCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG	883
QY	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	884	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG	943
QY	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAAGG	780
Db	944	CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG	1003
QY	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTGGCAGCTTTTGGG	840
Db	1004	GTCCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCTTCTTGGCAGCTTTTGGG	1063
QY	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1064	TGCCTGGTGATGGCTCTACCAGCCATTTGCATTGGGGCCATTGGAGCCTCCACAGACTGG	1123
QY	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1124	AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	1183
QY	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	1184	CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTTCTTTGGGCTTGGTGTCTGTT	1243
QY	1021	TCTGCTGCTGTTATGTCTCATCAGCAGATTCTTCCATCTTGTCTCAGCAAGTTCCATGTTTGCA	1080
Db	1244	TCTGCTGCTGTCTATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1303
QY	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1304	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1363
QY	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1364	ATGAGGATCACTGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1423
QY	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260

```

Db      1424 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG 1483
Qy      1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320
Db      1484 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1543
Qy      1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380
Db      1544 TTTGGACTTTTCCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1603
Qy      1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440
Db      1604 TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1663
Qy      1441 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500
Db      1664 ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCCTATCTAGCCAAGTAT 1723
Qy      1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560
Db      1724 CTATTTGAAAGTGGAACCTTGCCTCCAAAATTAGATATATTTGATGCTGTTGTCTCAAGG 1783
Qy      1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
Db      1784 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAAT 1843
Qy      1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
Db      1844 GAACTTGCACTTGTAAAGCCTCGACAGAGCCTAACCTCAGTCAACTTTCACCAATAAA 1903
Qy      1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
Db      1904 GAGGCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA 1963
Qy      1741 TGA 1743
Db      1964 TGA 1966

```

RESULT 4

US-09-911-077A-3

; Sequence 3, Application US/09911077A

; Publication No. US20030114399A1

; GENERAL INFORMATION:

; APPLICANT: BLAKELY, RANDY D.

; APPLICANT: APPARSUNDARAM, SUBRAMANIAM

; APPLICANT: FERGUSON, SHAWN

; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA

; FILE REFERENCE: VBLT:008US

; CURRENT APPLICATION NUMBER: US/09/911,077A

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 1743

; TYPE: DNA

```

;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1743)
US-09-911-077A-3

```

Query Match 78.9%; Score 1375; DB 10; Length 1743;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

QY	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGCCTTTCCATGTGGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG	60
QY	61	GTTGGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCATGGAAAACCACAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA	120
QY	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC	180
QY	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTTGGAGGAGGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	240
QY	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300
QY	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG	360
QY	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTC	420
Db	361	ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTC	420
QY	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
QY	481	ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
QY	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
QY	601	ATCAGCGTCCCCTTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGGCCCTGTACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	660
QY	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720

Qy 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 721 CTTGATAATTTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 780

Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 781 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTAAGTGTCTTCCTGGCAGCTTTCGGG 840

Qy 841 TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 841 TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG 900

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 901 AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC 960

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 961 CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT 1020

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA 1080
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1021 TCAGCTGCTGTCATGTCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT 1080

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1081 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC 1140

Qy 1141 ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGTGACGAAA 1200
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1141 ATGAGGATCACTGTGCTTGTGTTGCGAGCATCTGCAACAGCCATGGCTTGTGACGAA 1200

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTACATCGTTATCTTCCCCAG 1260
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1201 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTGTCTACATCATCATCTTCCACAG 1260

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1261 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1320

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1321 TTTGGACTATTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1381 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1440

Qy 1441 ACACCTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1441 ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1501 CTATTTGAAAGTGGAACCTTGCCCTCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG 1560

Db	181	ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGCCCTGTCTCACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGTCTGCTTCCCTGGCAGCTTTTGGG	840
Db	781	GTCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCCTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCTTCTTTGGGCTTGGTGTCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCTGAGCAAGTTCCATGTTTGCA	1080


```
; APPLICANT: APPARSUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 119040
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;     NAME/KEY: modified_base
;     LOCATION: (2347)..(90873)
;   OTHER INFORMATION: N = A, C, G or T/U
US-09-911-077A-19
```

QY	1110	AAATGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTGTTTGGAGC	1169
Db	30755	ACAGGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTGTTTGGAGC	30814
QY	1170	ATCTGCAACAGCCATGGCCTTGCTGACGAAAACGTGTGTATGGGCTCTGGTACCTCAGTTC	1229
Db	30815	ATCTGCAACAGCCATGGCCTTGCTGACGAAAACGTGTGTATGGGCTCTGGTACCTCAGTTC	30874
QY	1230	TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCCTTGTGTACTCTTTGTAAAGGGAAC	1289
Db	30875	TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCCTTGTGTACTCTTTGTAAAGGGAAC	30934
QY	1290	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	1349
Db	30935	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	30994
QY	1350	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	1409
Db	30995	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	31054
QY	1410	TGGTATATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATTCTTAAC	1469
Db	31055	TGGTATATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATTCTTAAC	31114
QY	1470	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	1529
Db	31115	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	31174
QY	1530	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	1589
Db	31175	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	31234
QY	1590	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	1649
Db	31235	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	31294

Qy 1650 CATGACCCTCAGCTCAACTTTACCAATAAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 1709
 |||
 Db 31295 CATGACCCTCAGCTCAACTTTACCAATAAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 31354

Qy 1710 AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 1743
 |||
 Db 31355 AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 31388

RESULT 7

US-09-911-077A-14
 ; Sequence 14, Application US/09911077A
 ; Publication No. US20030114399A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BLAKELY, RANDY D.
 ; APPLICANT: APPARSUNDARAM, SUBRAMANIAM
 ; APPLICANT: FERGUSON, SHAWN
 ; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
 ; FILE REFERENCE: VBLT:008US
 ; CURRENT APPLICATION NUMBER: US/09/911,077A
 ; CURRENT FILING DATE: 2001-07-23
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 142299
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Primer
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (1305)..(127835)
 ; OTHER INFORMATION: N = A, C, G or T/U
 US-09-911-077A-14

Query Match 36.2%; Score 630.8; DB 10; Length 142299;
 Best Local Similarity 99.7%; Pred. No. 1.4e-178;
 Matches 632; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1110 AAATGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTGTTGGAGC 1169
 |||
 Db 94673 ACAGGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTGTTGGAGC 94732

Qy 1170 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 1229
 |||
 Db 94733 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 94792

Qy 1230 TGACCTTGTTTACATCGTTATCTTCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC 1289
 |||
 Db 94793 TGACCTTGTTTACATCGTTATCTTCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC 94852

Qy 1290 CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG 1349
 |||
 Db 94853 CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG 94912

QY 1350 GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA 1409
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 94913 GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA 94972

QY 1410 TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAAC 1469
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 94973 TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAAC 95032

QY 1470 CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA 1529
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 95033 CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA 95092

QY 1530 ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT 1589
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 95093 ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT 95152

QY 1590 TCTTGTCAAAAATGAAAATATTAAATTAGATGAACCTTGCACTTGTGAAGCCACGACAGAG 1649
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 95153 TCTTGTCAAAAATGAAAATATTAAATTAGATGAACCTTGCACTTGTGAAGCCACGACAGAG 95212

QY 1650 CATGACCCTCAGCTCAACTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 1709
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 95213 CATGACCCTCAGCTCAACTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 95272

QY 1710 AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 1743
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 95273 AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 95306

RESULT 8

US-10-241-784-1

; Sequence 1, Application US/10241784

; Publication No. US20040048261A1

; GENERAL INFORMATION:

; APPLICANT: Bayer Corporation

; TITLE OF INVENTION: Invertebrate Choline Transporter Nucleic Acid,
 Polypeptides and Uses

; TITLE OF INVENTION: Thereof

; FILE REFERENCE: MO7218

; CURRENT APPLICATION NUMBER: US/10/241,784

; CURRENT FILING DATE: 2002-09-11

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1833

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1833)

; OTHER INFORMATION:

US-10-241-784-1

Query Match 21.6%; Score 376.6; DB 13; Length 1833;
 Best Local Similarity 56.0%; Pred. No. 9.9e-103;
 Matches 868; Conservative 0; Mismatches 624; Indels 57; Gaps 6;

QY 8 TCCATGTGGAAGGACTGATAGCTATCATCGTGTTCACCTTCTAATTTTGCTGGTTGGAA 67
| | | | | | | | | | | | | | | | | | | | |
Db 5 TCAATATCGCTGGCGTGGTGAGCATCGTGCTCTTCTACCTCCTGATCCTGGTCGTTGGCA 64

QY 68 TATGGGCTGCCTGGAGAACC AAAACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATCA 127
| | | | | | | | | | | | | | | | | | | | |
Db 65 TT TGGGCCGGTCGCAAGAAGCAGTCCGGCAATGATT CGGAGGAG-----GAGGTCA 115

QY 128 TAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATT TACCATGACAGCTACCTGGG 187
| | | | | | | | | | | | | | | | | | | | |
Db 116 TGCTGGCCGGACGCTCCATCGGCCTCTTCGTGGGCATCTTCACCATGACGGCCACCTGGG 175

QY 188 TCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTAG 247
| | | | | | | | | | | | | | | | | | | | |
Db 176 TGGGTGGCGGCTACATCAACGGCACGGCGGAGGCTATATAACACATCGGGT-----CTGG 229

QY 248 CTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTG 307
| | | | | | | | | | | | | | | | | | | | |
Db 230 TGTGGT GCCAGGCTCCATTTGGATACGCTCTAAGCTTGGTATTTGGTGGCATCTTCTTTG 289

QY 308 CAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATG 367
| | | | | | | | | | | | | | | | | | | | |
Db 290 CCAATCCCATGCGCAAGCAGGGTTACATCACCATGTTGGATCCGTTGCAGGATTCCTTTG 349

QY 368 GAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTCTGGGCTG 427
| | | | | | | | | | | | | | | | | | | | |
Db 350 GTGAGCGGATGGGAGGATTGCTCTTCTGCCCGCTCTATGCGGTGAGGTCTTTTGGGCAG 409

QY 428 CAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCACA 487
| | | | | | | | | | | | | | | | | | | | |
Db 410 CCGGCATCCTGGCTGCACTTGGCGCCA CTCTATCGGTGATCATCGACATGGATCACC GCA 469

QY 488 TTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATT 547
| | | | | | | | | | | | | | | | | | | | |
Db 470 CCTCGGTGATCCTGTCTCTG CATCGCCATCTTCTACACACTGTTCCGTGGACTGTACT 529

QY 548 CTGTGGCCTACACTGATGTCGTT CAGCTCTTTTGCATTTTTGTAGGGCTGTGGATCAGCG 607
| | | | | | | | | | | | | | | | | | | | |
Db 530 CCGTGGCGTATACGGACGTGATCCAGTTGTTCTGCATCTTCATCGGTCTGTGGATGTGCA 589

QY 608 TCCCCCTTTGCATTGTACATCCTG CAGTCGCAGACATCGGGTTCACTGCTGTGCATGCCA 667
| | | | | | | | | | | | | | | | | | | | |
Db 590 TTCCCTTCGCCTGGAGCAACGAGCACGTGGGCAGCCTGAGTGACCTGGAGGTGGAT---- 645

QY 668 AATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGGCTTGATA 727
| | | | | | | | | | | | | | | | | | | | |
Db 646 -----TGGATTGGGCACGTGGAGCCTAAAAAGCATTGGCTGTACATAGACT 691

QY 728 GTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTT CAGAGGGTTCTCT 787
| | | | | | | | | | | | | | | | | | | | |
Db 692 ACGGCTTGCTGCTCGTCTTTGGTGGCATTCCCTGGCAGGTCTACTTCCAGCGGCAAAAC- 750

QY 788 CTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTT CGGGTGCCTGG 847
| | | | | | | | | | | | | | | | | | | | |
Db 751 -----GGCAGGAAGGGCCAGCTTCTGCCTATGTTGCAGCCGCCGGATGCATTT 799

QY 848 TGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAACCAGA 907


```
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1739)
US-09-911-077A-7
```

```
Query Match          20.9%; Score 363.8; DB 10; Length 1985;
Best Local Similarity 55.1%; Pred. No. 8.1e-99;
Matches 862; Conservative 0; Mismatches 637; Indels 66; Gaps 5;
```

```
Qy      19  GGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTGGTTGGAATATGGGCTGCC 78
        || | | | | | | | | | | | | | | | | | | | | | | | |
Db      24  GGTATCGTGGCCATTGTGTTCTTCTACGTGCTCATTCTTGTCGTTGGAATATGGGCGGGT 83

Qy      79  TGGAGAACCAAAA-----ACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATC 126
        | | | | | | | | | | | | | | | | | | | | | | | |
Db      84  AGAAAATCGAAAAGTTCAAAAGAGCTTGAATCAGAAGCCGGCGCGGCGACGGAAGAGGTG 143

Qy     127  ATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCTACCTGG 186
        || | | | | | | | | | | | | | | | | | | | | | |
Db     144  ATGTTAGCTGGGAGAAACATCGGAACTCTTGTCGGAATTTTCACAATGACTGCCACGTGG 203

Qy     187  GTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTA 246
        || || | | | | | | | | | | | | | | | | | | | |
Db     204  GTTGGCGGTGCTTATATCAATGGAACCGCCGAGGCTCTGTATAATGGAGGT-----CTC 257

Qy     247  GCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTT 306
        | | | | | | | | | | | | | | | | | | | | | | | |
Db     258  CTTGGATGTCAGGCTCCAGTTGGATATGCAATTTCCCTTGTTATGGGAGGACTACTTTTC 317

Qy     307  GCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTAT 366
        || || | | | | | | | | | | | | | | | | | | | |
Db     318  GCAAAGAAAATGCGAGAAGAAGGATATATTACAATGCTCGATCCTTTTCAGCACAAATAT 377

Qy     367  GGAAAACGCATGGGCGGACTCCTGTTTATTTCCTGCACTGATGGGAGAAATGTTCTGGGCT 426
        || || || | | | | | | | | | | | | | | | | | | |
Db     378  GGCCAACGAATCGGTGGCTTGATGTATGTTCCAGCACTTCTTGGTGAAACATTCTGGACA 437

Qy     427  GCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCAC 486
        || || | | | | | | | | | | | | | | | | | | | |
Db     438  GCAGCCATTCTTTTCGGCACTTGGTGCAACACTGTGCGTAATTCTTGAATCGACATGAAT 497

Qy     487  ATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTAT 546
        || || | | | | | | | | | | | | | | | | | | | |
Db     498  GCATCAGTGACCCTGTGCGGCTGTATTGCCGTATTCTACACATTACCGGTGGATACTAT 557

Qy     547  TCTGTGGCCTACACTGATGTGCTTCAGCTCTTTTGCATTTTGTAGGGCTGTGGATCAGC 606
        | | | | | | | | | | | | | | | | | | | | | | | |
Db     558  GCAGTCGCGTACACTGACGTGCTTCAACTATTTTGCATTTTTCGTCGGTTTGTGGGTTTGC 617
```

QY	607	GTCCCTTTGCATTGTCACATCCTGCAGTGCAGACATCGGGTTCAC TGCTGTGCATGCC	666
Db	618	GTGCCGGCGGCTATGGTGCATGATGGTGCGAAGGATATTTCCAGGAATGCAG-----	669
QY	667	AAATACCAAAAGCCGTGGCTGGGAAGCTGTTGACTCATCTGAAGTCTACTCTTGGCTTGAT	726
Db	670	-----GCGACTGGATTGGAGAGATTGGAGGATTCAAAGAAACATCTCTCTGGATTGAT	722
QY	727	AGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTT CAGAGGGTTCTC	786
Db	723	TGCATGCTTCTCCTTGCTTTGGAGGAATTCCATGGCAAGTGTACTTCCAAGAGTTCTC	782
QY	787	TCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCCTGGCAGCTTTCGGGTGCCTG	846
Db	783	TCCTCAAAAAGTGTCTCATGGAGCACAGACGTTGTCTGTTTGTGGCGGGCGTCGGATGCATT	842
QY	847	GTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAACCAG	906
Db	843	CTCATGGCGATTCCACCAGCGTTGATCGGTGCAATTGCCAGGAACACAGACTGGAGAATG	902
QY	907	ACTGCATATGGGCTTCC-----AGATCCCAAGACTACAGAAGAGGCA	948
Db	903	ACTGATTATCCCCATGGAACAATGGAAGTAAGGTCGAATCGATTCCACCGGATAAGAGA	962
QY	949	GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT	1008
Db	963	AACATGGTGGTCCCGTTGGTATTCCAGTATCTTACGCCAAGATGGGTGCCTTTATTGGA	1022
QY	1009	CTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGT CAGCAAGT	1068
Db	1023	CTCGGCGCAGTGTCTGGCTGCTGTAATGTCATCTGCAGATT CATCTGTACTATCAGCAGCA	1082
QY	1069	TCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAA	1128
Db	1083	TCAATGTTTGTCTACAACATCTGGAAGCTCACAATTGCCCCCTCACGCGTCTGAAAAAGAA	1142
QY	1129	ATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCC	1188
Db	1143	GTGATAATTGTGATGAGAATAGCCATCATCTGTGTTGGTATCATGGCAACCATCATGGCA	1202
QY	1189	TTGCTGACGAAAAGCTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTT	1248
Db	1203	CTTACCATTCAATCCATCTATGGGCTTTGGTATCTTTGTGCAGATTGGTCTACGTCATA	1262
QY	1249	ATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTG	1308
Db	1263	CTCTTCCCTCAACTATTATGTGTTGTATATATGCCACGTAGCAATACGTATGGCTCATTG	1322
QY	1309	GCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTT	1368
Db	1323	GCTGGCTATGCAGTCGGTCTTGTGCTCCGTTTGATTGGAGGCGAGCCACTTGTATCGCTG	1382
QY	1369	CAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAA	1428
Db	1383	CCAGCGTCTCTCCATTATCCAATGTATACGGATGGGG-----TACAGTAT	1427

Qy 1429 TTTCCATTTAAAAACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTAT 1488
 || ||||| | ||| ||| ||| | | ||| | ||| |
 Db 1428 TTCCCATTCAGGACAACCTGCTATGTTATCTTCAATGGCTACTATCTACATTGTATCAATA 1487
 Qy 1489 CTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCT 1548
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1488 CAATCGGAGAAGCTGTTCAAATCGGGACGTTTGTCTCCGGAGTGGGACGTAATGGGTTGT 1547
 Qy 1549 GTTGT 1553
 || ||
 Db 1548 GTAGT 1552

RESULT 10

US-09-974-300-501

; Sequence 501, Application US/09974300

; Patent No. US20020146721A1

; GENERAL INFORMATION:

; APPLICANT: Berka, Randy M.

; APPLICANT: Clausen, Ib Groth

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene

; TITLE OF INVENTION: Expression

; FILE REFERENCE: 10085.500-US

; CURRENT APPLICATION NUMBER: US/09/974,300

; CURRENT FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 09/680,598

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/279,526

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 8481

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 501

; LENGTH: 1461

; TYPE: DNA

; ORGANISM: Bacillus licheniformis

US-09-974-300-501

Query Match 13.9%; Score 242.6; DB 9; Length 1461;
 Best Local Similarity 52.4%; Pred. No. 4.3e-62;
 Matches 663; Conservative 0; Mismatches 554; Indels 48; Gaps 4;

Qy 151 TTATTGGTTGGTGGATTACCATGACAGCTACCTGGGTCGGAGGAGGGTATATCAATGGC 210
 || || || || | || ||| | || || ||||| || || ||||| || ||
 Db 125 TTTTTCGTCGGAATGGTGACGATGGCCGCAACATGGGTCGGCGGCGGATATATTAACGGA 184
 Qy 211 ACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTAGCTTGGGCTCAGGCACCAATTGGA 270
 || || ||| | ||| | | |||| | |||| || || || ||
 Db 185 ACGCCGAATCGACTTACA-----GCGACGGCCTCATCTGGGCCCCAAGCGCCTTGGGGC 238
 Qy 271 TATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGCAAAACCTATGCGTTCAAAGGGG 330
 || | | || ||||| | || || | |||| || || | |||| |
 Db 239 TACGCATTGAGCCTGATTATCGGCGGTATTTTCTTCGCCAGAAAAATGCGCCGTCATCAA 298
 Qy 331 TATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGGAAAACGCATGGGCGGACTCCTG 390
 | | ||||| | || || | || || || || || || || || ||
 Db 299 TTTATGACCATATCGATCCTCTCGAACAGCGCTTCGGTAAGCGGATGGCCGGCGTTTAA 358

Qy 391 TTTATTCTGCACTGATGGGAGAAATGTTCTGGGCTGCAGCAATTTTCTCTGCTTTGGGA 450
 | | | | | | | | | | | | | | | | | | | | | |
 Db 359 TATATACCGGCGCTGTTAGGAGAATTGTTTTGGAGCGCCGCGATCTTAACGGCATTGGGC 418

 Qy 451 GCCACCATCAGCGTGATCATCGATGTGGATATGCACATTTCTGTCATCATCTCTGCACTC 510
 | | | | | | | | | | | | | | | | | | | | | |
 Db 419 ACGACTTTCGGAATGATTCTGAATATCGATTTCCAAACGTCGATTATTCTTTCGGCGATG 478

 Qy 511 ATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTCTGTGGCCTACACTGATGTCGTT 570
 | | | | | | | | | | | | | | | | | | | | | |
 Db 479 ATCGCCATCGCTTATACGGTGGCCGGAGGCATGTGGGCAGTTGCTTTACAGATGTCTTT 538

 Qy 571 CAGCTCTTTTGCATTTTGTAGGGCTGTGGATCAGCGTCCCCTTTGCATTGTACATCCT 630
 | | | | | | | | | | | | | | | | | | | | | |
 Db 539 CAAATGATTGTCATTTTGTCTCGGGCTGTTTTTAGTCGTCCCATTTGTATTGTGCAATGTC 598

 Qy 631 GCAG-----TCGCAGACATCGGGTTCACTGCTGTGCATGCCAA 668
 | | | | | | | | | | | | | | | | | | | | | |
 Db 599 GGTGCTCTCGATAGCGTATGGGCAAATTACAGGCACGATTTCCGCAGCAGCGCCAATCTG 658

 Qy 669 ATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCT-----TGGCTT 723
 | | | | | | | | | | | | | | | | | | | | | |
 Db 659 CTTCCGCCGTTGGACGGCTGGAAAAACCCGGATTGGGGAACCTGTTTTGGAAGTGGTGG 718

 Qy 724 GATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAAGGGT 783
 | | | | | | | | | | | | | | | | | | | | | |
 Db 719 GATAATGCGCTCCTCTTAATTTTCGGAGGTATCGCATGGCAGGTGTACTTTTCAAGCGT 778

 Qy 784 CTCTCTTCTTCTCAGCCACCTATGCTCAAGTGTCTTCTTCTTGGCAGCTTTTCCGGTGC 843
 | | | | | | | | | | | | | | | | | | | | | |
 Db 779 CTTTCGGCAAAATCGGAAAGCGCCGCCATGTGGCAGTCGATAATTGCCGGAGTGATCTGC 838

 Qy 844 CTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAAC 903
 | | | | | | | | | | | | | | | | | | | | | |
 Db 839 ATCATTGCCGCCATTCCGTGCGTAATCATCGGAGCTGCCGGAACAGTACCGATTGGAGC 898

 Qy 904 CAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTACCA 963
 | | | | | | | | | | | | | | | | | | | | | |
 Db 899 CTGTTCCGAGCGAGCGCTCCGGATAACCCGGCG-----ATGATTTTGCCG 943

 Qy 964 ATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTTTCT 1023
 | | | | | | | | | | | | | | | | | | | | | |
 Db 944 CAAACGCTTGCGTATTTGACGCCAGGAATCATCGCAGGCCTCGGCTTGGGTGCAATCGCA 1003

 Qy 1024 GCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCACGG 1083
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1004 GCAGCCGTCTATGTCAGCATGGACTCATCGATTCTATCGGCATCATCAATGGCCGCATGG 1063

 Qy 1084 AACATCTACCAGCTTTTCTTCCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTTATG 1143
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1064 AATATTTACCGTCCGCTCATCAAGCCGAAGGCCACCCAAAACAGCTGCAAAAAGTCGTC 1123

 Qy 1144 CGAATCACAGTGTGTTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAACT 1203
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1124 AAACGCTCAATCATTTTGTTCGGCGCGGAGCAGCGGTATCGCGCTGAATGTCAAAGC 1183

 Qy 1204 GTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTACATCGTTATCTTCCCCAGCTG 1263

```

      || |||   | ||||| |   ||| || | ||||| | || | || ||||| |
Db      1184 GTTTATACTTTATGGTATTTGGCTTCGGATTAGTTTATTGCATTCTTTTCCCCAGTTA 1243
Qy      1264 CTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTTTCT 1323
      | ||||| ||| || | || | || | || | || | || |
Db      1244 ACAATGGCCCTCTTTTATAAAAGAGCAAATCTTTACGGGTCGATTGCTGGATTGTCAGTT 1303
Qy      1324 GGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATCTTC 1383
      | || | ||||| |   || || || || |   || || || | |
Db      1304 GCAGTCATTCTGAGGCTCGGCGGTGGTGAACCCGCATTCGGCATTCGCCCGCTTCTGCCG 1363
Qy      1384 TACCC 1388
      || ||
Db      1364 TATCC 1368

```

RESULT 11

US-09-911-077A-19/c

; Sequence 19, Application US/09911077A

; Publication No. US20030114399A1

; GENERAL INFORMATION:

; APPLICANT: BLAKELY, RANDY D.

; APPLICANT: APPARSUNDARAM, SUBRAMANIAM

; APPLICANT: FERGUSON, SHAWN

; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA

; FILE REFERENCE: VBLT:008US

; CURRENT APPLICATION NUMBER: US/09/911,077A

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 19

; LENGTH: 119040

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: modified_base

; LOCATION: (2347)..(90873)

; OTHER INFORMATION: N = A, C, G or T/U

US-09-911-077A-19

```

Query Match          10.4%; Score 180.8; DB 10; Length 119040;
Best Local Similarity 98.9%; Pred. No. 6.3e-42;
Matches 182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      |||||
Db      94584 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 94525
Qy      61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      |||||
Db      94524 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 94465
Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      |||||
Db      94464 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGGT 94405
Qy      181 ACCT 184

```

Db || |
 94404 ACGT 94401

RESULT 12

US-09-911-077A-14/c
; Sequence 14, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARSUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 142299
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (1305)..(127835)
; OTHER INFORMATION: N = A, C, G or T/U
US-09-911-077A-14

Query Match 10.4%; Score 180.8; DB 10; Length 142299;
Best Local Similarity 98.9%; Pred. No. 7.2e-42;
Matches 182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
 |||
Db 64222 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 64163

QY 61 GTTGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
 |||
Db 64162 GTTGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 64103

QY 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
 |||
Db 64102 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGGT 64043

QY 181 ACCT 184
 |||
Db 64042 ACGT 64039

RESULT 13

US-09-864-761-1838
; Sequence 1838, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

```

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1838
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009963.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2

```

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.97
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
 US-09-864-761-1838

Query Match 8.9%; Score 155; DB 9; Length 455;
 Best Local Similarity 100.0%; Pred. No. 7.5e-36;
 Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 741 GATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGGGTTCTCTCTTCTTCCTCAGC 800
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 266 GATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGGGTTCTCTCTTCTTCCTCAGC 325
 Qy 801 CACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGCGGTGCCTGGTGATGGCCATCCC 860
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 326 CACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGCGGTGCCTGGTGATGGCCATCCC 385
 Qy 861 AGCCATACTCATTGGGGCCATTGGAGCATCAACAG 895
 ||||||||||||||||||||||||||||||||
 Db 386 AGCCATACTCATTGGGGCCATTGGAGCATCAACAG 420

RESULT 14

US-10-027-632-120553/c
 ; Sequence 120553, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 120553
 ; LENGTH: 943
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-120553

Query Match 6.8%; Score 118.6; DB 13; Length 943;
Best Local Similarity 99.2%; Pred. No. 1.5e-24;
Matches 118; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      176 CAGCTACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAG 235
          |||
Db      589 CAGCTACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAG 530

Qy      236 GTTATGGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGT 294
          |||
Db      529 GTTATGGCCTAGCTTGGGCTCAGGCACCARTTGGATATTCTCTTAGTCTGATTTTAGGT 471
```

RESULT 15

US-10-027-632-120553/c

; Sequence 120553, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 120553

; LENGTH: 943

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-120553

Query Match 6.8%; Score 118.6; DB 16; Length 943;
Best Local Similarity 99.2%; Pred. No. 1.5e-24;
Matches 118; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      176 CAGCTACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAG 235
          |||
Db      589 CAGCTACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAG 530

Qy      236 GTTATGGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGT 294
          |||
Db      529 GTTATGGCCTAGCTTGGGCTCAGGCACCARTTGGATATTCTCTTAGTCTGATTTTAGGT 471
```

Search completed: September 29, 2004, 05:05:29
Job time : 1195 secs